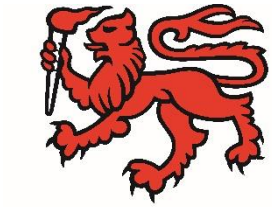


Mapping species composition and structure in wet eucalypt forest using multi-source remote sensing data

By

Bechu Kumar Vinwar Yadav

M. Sc. Forestry, Master's Degree in Sociology, M. Tech. (Remote Sensing and GIS)
School of Technology, Environments, and Design
College of Sciences and Engineering

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy in Geomatics Engineering
University of Tasmania
October 2019

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Bechu Kumar Vinwar Yadav
11 November 2019

Authority of Access

Acknowledgements

First, I would like to express my sincere gratitude to my supervisory panel, Associate Professor Arko Lucieer, Associate Professor Greg Jordan and Dr Sue Baker for their contribution, invaluable guidance, and encouragement throughout my PhD study. I am deeply grateful for the opportunity to work under the supervision of academics with both knowledge and passion for the subject area. My special thanks go to my primary supervisor Associate Professor Arko Lucieer for accepting me as his PhD student and for offering his kind advice, encouragement and tireless guidance to move forward leading me to this day of completing my research.

I would like to acknowledge the University of Tasmania, Australia for providing me with a Graduate Research Scholarship and the Australian Research Council for providing me with a top-up stipend to carry out my PhD research work. My sincere thanks go to the staff at the Discipline of Geography and Spatial Sciences for their support; including special thanks to Senior Lecturer Dr Jon Osborn and Mr Robert Anders. I am grateful to Associate Professor Greg Jordan for providing me with a small grant to complete my study. Thanks to Dr Tim Wardlaw for providing me with the detailed field data for the research sample plot.

I am thankful to the Chair of Examiners, Distinguished Professor Jamie Kirkpatrick, and research committee members- Professor Matt King, Professor Jason Byrne, and Graduate Research Coordinator Dr Dave Kendal, for their timely consideration and recommendation. I would like to acknowledge my thesis examiners- Professor Simon Jones and Associate Professor Hooman Latifi for their thorough review and genuine suggestions.

I really appreciate my beloved family members and friends. Special thanks go to my colleagues Dr Steve Harvin, Sachit Rajbhandari, Irfan Iqbal, and Romain Louvet for sharing their knowledge in times of need. Thanks also go to Dr Ram Asheshwar Mandal, Kamlesh Yadav, and Dr Buddi Poudel for their encouragement. Many thanks to my younger brother Shreedeo and Uncle Ramjatan Vinwar.

Last, but not least, thank you to my son Ankit and daughter Aayusha. I am so grateful to have such lovely children who have been very understanding and adapted to these new circumstances without my guidance and support. Most importantly, my PhD study would never have been possible without the support and encouragement of my lovely wife Renu. Thank you, my love! The sincerest appreciation goes to my parents for all their love and support over the years. Thank you, my mom, Jago Devi and dad Ramanand Vinwar Yadav for giving me a positive upbringing and pushing me for education. I would like to dedicate this thesis to my parents. Thanks, God!

Abstract

Tasmanian wet eucalypt forests are internationally important for wood and paper production, carbon storage and biodiversity conservation. These forests contain tall eucalypts over dense understories of rainforest and wet sclerophyll species. This research was motivated by a need for tools to replace costly aerial photo interpretation (PI-type) mapping for describing forest species composition and stand structure. Overall, I aimed to develop approaches for assessing and mapping tree species distribution and forest structure of wet eucalypt forest in a 5 km by 5 km area of the Warra Supersite, Tasmania, using multi-source remote sensing data.

My first study used an airborne LiDAR-derived canopy height model (CHM) and hyperspectral imagery to classify up to five dominant tree species of the forest. I used random forest classifiers on objects generated using data segmentation under a range of scenarios. Fused CHM and Minimum Noise Fraction (MNF) datasets yielded the highest segmentation accuracy (88.71%). The fusion of hyperspectral imagery, CHM and vegetation indices produced the best classifiers (overall accuracy (OA) of 66.7%) followed by the fused dataset of MNF and CHM (OA = 66.0%). Hyperspectral imagery alone provided the lowest classification accuracy (OA = 59.0%). Accuracy for the dominant canopy species (*Eucalyptus obliqua*) was 90.86% for four vegetation classes and 86.11% for five classes. Classification accuracies for the important understory species, *Dicksonia antarctica*, were also high under the best models (~84%). Accuracies for other species were low. Thus, fused hyperspectral and LiDAR data were robust and capable of spatially discriminating several important forest species.

My second study utilised LiDAR-derived topographic attributes and mapped geological strata to develop a model for predicting three understory layers of the forest (≥2 to ≤10 m, >10 to ≤30 m and >30 to ≤50 m as proxies for the lower, middle and upper layers, respectively) using five different spatial resolutions using random forest regression. Overall, the 30 m resolution provided the best model for predicting understory layers compared to 1 m, 5 m, 10 m and 20 m resolutions. The predictive power for the upper layer was greatest ($R^2 = 0.82$), followed by the lower layer and the middle layer. Geology had the highest variable importance score for 5 m, 10 m, 20 m and 30 m resolutions, whereas terrain position index had the highest variable importance score for 1 m resolution. This research demonstrated that LiDAR-derived

topographic attributes and geology data could be used to predict the understory vegetation structure.

My third study developed robust and cost-effective approaches for predicting the densities of vertical structural layers of the forest based on multispectral satellite data and simulated operational LiDAR datasets. I assessed the robustness of forest structure models based on thirteen schemes of derivatives (vegetation indices, texture features, and topographic attributes) from three different data sources (Airborne LiDAR downscaled to operational density, WorldView-3 and Landsat-8 (OLI)) at spatial resolutions (1.60 m, 7.5 m and 30 m). Models for the upper and middle layers were better than those for the lower layer. The 30 m Landsat-8 data provided the best results for all three-pixel sizes (R^2 values ranged 0.15 to 0.65). Fused data from Landsat-8 and the simulated low-density LiDAR showed modest accuracy for predicting the density of three vertical layers and could be adopted by forest managers and planners. The WorldView-3 data of 1.6 m pixel size did not produce useful models.

In conclusion, the fusion of remote sensing datasets may help assess and map woody plant species composition and structure of wet eucalypt forests with opportunities to replace the traditional, subjective and time-consuming mapping technique of aerial photo interpretation. My results highlight the potential of freely available Landsat-8 (OLI) and operational LiDAR data, and random forest machine learning techniques for predicting and mapping forest species and vertical structural layers of wet eucalypt forests. This thesis addressed data complexities, including multidimensionality and nonlinearity in multi-source data, and provided a robust approach for the assessment of wet eucalypt forest composition and structure.

List of Abbreviations

%IncMSE	Percent Increase in Mean Square Error
ALS	Airborne Laser Scanning
ARN	Aggregated Retention
BA	Balance Accuracy
CART	Classification And Regression Trees
CBS	Clearfelling, Burning and Aerial Sowing (clearcutting)
CHM	Canopy Height Model
CI	Convergence Index
CPA	Crown Projection Area
DEM	Digital Elevation Model
DN	Digital Number
DSM	Digital Surface Model
DTM	Digital Terrain Model
FLAASH	Fast Line-of-Sight Atmospheric Analysis of Hypercubes
FWHM	Full-Width-at-Half-Maximum
GARI	Green Atmospherically Resistant Index
GDA	Geocentric Datum of Australia
GIS	Geographic Information System
GLCM	Grey-Level Co-Occurrence Matrix
GNDVI	Green Normalized Difference Vegetation Index
GPS	Global Positioning System
GRVI	Green Ratio Vegetation Index
HSI	Hyperspectral Imagery
ICA	Independent Component Analysis
IPVI	Infrared Percentage Vegetation Index
LiDAR	Light Detection And Ranging
LTER	Long Term Ecological Research
MDA	Mean Decrease in Accuracy
MDG	Mean Decrease in Gini
MGA	Map Grid of Australia
MNF	Minimum Noise Fraction
MNLI	Modified Non-Linear Index
MRMSE	Mean of Root Mean Square Errors

MRS	Multiresolution Segmentation
MSAVI	Modified Soil Adjusted Vegetation Index
MSR	Modified Simple Ratio
NBR	Normalized Burn Ratio
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-Infrared
NLI	Non-Linear Index
OA	Overall Accuracy
OBIA	Object-Based Image Analysis
OLI	Operational Land Imager
OOB	Out Of Bag
OSAVI	Optimized Soil Adjusted Vegetation Index
PI	Photograph Interpretation
QUAC	Quick Atmospheric Correction
RDVI	Renormalized Difference Vegetation Index
RF	Random Forest
SAGA	System for Automated Geoscientific Analyses
SATVI	Soil-Adjusted Total Vegetation Index
Se	Sensitivity
SIMI	Shortwave Infrared Soil Moisture Index
Sp	Specificity
SPI	Stream Power Index
SRVI	Simple Ratio Vegetation Index
SVM	Support Vector Machine
SWCI	Surface Water Capacity Index
SWI	Saga Wetness Index
SWIR	Shortwave Infrared
TERN	Terrestrial Ecosystem Research Network
TPI	Terrain Position Index
TRI	Terrain Ruggedness Index
USGS	United States Geological Survey
VI	Variable Importance
VIF	Variable Inflation Factor
VNIR	Visible Near-Infrared
VR	Variable Retention

Table of Contents

Chapter 1

Introduction

1.1 Research background

1.1.1 Wet eucalypt forests and their management approaches

In Tasmania, wet eucalypt forests are highly productive sources of hardwood for timber and paper (Baker 2006; Scott et al. 2015), and the most widespread commercial native forests, covering 12% (804, 000 ha) of Tasmania's land area (Forest Practices Authority 2017). They are important for wood and pulp production (e.g. Elliott et al. 2008; Yee 2005), carbon storage and biodiversity (e.g. Dean et al. 2012; Hickey 1994a; Moroni et al. 2010). These wet eucalypt forests are also important as a model system for forest succession (Jackson 1968). They are dominated by a tall (typically 30-90 m) open eucalypt canopy (Attiwill et al. 2014), usually made up of *Eucalyptus obliqua* and/or *E. regnans* (Forest Practices Authority 2017; Scanlan et al. 2010). These forests occur in areas of high rainfall and are a subset of tall wet eucalypt forest types extending from Queensland to southern Tasmania, and to the southwest of Western Australia (Hickey et al. 2006).

These forests develop complex understories due to the open nature of the eucalypt crowns (Hickey et al. 2006) and are extensively found in areas of moderate to high rainfall and on varieties of soils (Attiwill et al. 2014). For the purpose of this thesis, I define understory to include all vegetation underneath the eucalypt canopy. Note, for some of these forests it is also possible to consider the eucalypts as emergent trees over a canopy of other species (Tng et al. 2012). That definition was avoided because it might make it difficult to compare structures across the full range of these wet eucalypt forests.

The understories of Tasmania's wet eucalypt forests broadly fall into two categories that are used to define two main forest types (mixed forest and wet sclerophyll forest). The understories of wet sclerophyll forests are dominated by broad-leaved, early successional trees and shrubs, whereas mixed forests have understories dominated by rainforest trees and woody shrub species. The vascular and non-vascular flora in the understories of these mixed forests are similar to those of the rainforests in this region (Forestry Tasmania 2009a). In reality, the wet sclerophyll and mixed forest types intergrade with each other and are usually considered to represent successional stages, post-fire (Jackson 1968; van Galen et al. 2018; Wood and Bowman 2012).

The management of wet eucalypt forests has long been the subject of public debate over balancing wilderness values, the conservation of biodiversity and wood

production (Baker 2006). Anthropogenic disturbances change species composition and forest structure resulting in habitat loss and species community declines over a wide range of geographical areas (Lindberg et al. 2015). This emphasizes the demand for detailed information on forest biological diversity at different spatial scales to detect and preserve habitat for varieties of flora and fauna (Lesak et al. 2011). Since the 1960s, silviculture in wet eucalypt forests has mainly involved clearfelling followed by high-intensity burning and aerial sowing (CBS). Prior to this practice, selective logging was common (Hickey et al. 2001). CBS is the most commonly prescribed system for harvesting wet eucalypt forests that meets all the key silvicultural considerations (Forestry Tasmania 2009a; Hickey et al. 2006). Clearfelling is generally applied for rotations of about 45-90 years in an average of 50 ha coupes (Hickey et al. 2006; Hickey et al. 1999). Nonetheless, wet eucalypt species require a disturbance to regenerate and the most common disturbance in Tasmania is wildfire that opens the canopy for seed-fall and prepare a mineral seed bed (Turner et al. 2009). For this reason, regeneration burning is standard practice following the harvesting of wet eucalypt forests and is important for the successful establishment of a new eucalypt cohort (Neyland et al. 2009). Although intense wildfires kill many plants, understory species establish quicker, and some plants regenerate from the existing soil-stored seeds (Baker et al. 2013), wind-dispersed seed, seed dispersed by birds and coppicing (Tabor et al. 2007). Timber logging and removal of large fuels from the clearfelled coupes, and disturbance to top soil and organic matter have an impact on regeneration.

Of the alternatives to clearfelling, variable retention (VR) techniques are considered to be much 'closer to nature' for wet eucalypt forests in south-eastern Australia than continuous-cover techniques that are practiced in Central European forests (Hickey et al. 2015). These techniques retain a portion of the original forest stand throughout the following rotation, either as small patches (aggregates) or as dispersed trees (Scott et al. 2013). VR approaches to sustainable forest management greatly improve the conservation of biodiversity and maintenance of key ecosystem processes in forests (Hickey et al. 2015; Lindenmayer et al. 2012). In wet eucalypt forests, an aggregated retention (ARN) form of VR has been suggested to be the most suitable alternative to clearfelling (Forestry Tasmania 2009a; Neyland et al. 2012). The ARN system was first applied in 2004-2005 and more than one hundred ARN coups have been harvested across Tasmania since then (Baker et al. 2017; Hindrum et al. 2012; Scott et al. 2013).

1.1.2 Past and current practices in mapping Tasmanian forests

In Tasmania and Australia, forest mapping started with hand-drawn sketch maps, but aerial photography was introduced in 1933 to produce topographic and forest type maps that provide significantly more accurate maps and estimates of the wood resources. By the early 1950s, the skills of the photo-interpreters with stereo-paired aerial photos had developed to predict stand height and density of canopy strata in combination with field-based inventory as validation. The codes of photo-interpreted types (PI types) have been modified over the years to expand the objectives of producing forest type and topographic maps and increased the efficacy of forest inventory (Elliott et al. 2008). PI type maps have been a key tool for specific purposes, for example, marking forest boundaries, mapping old-growth forests, site productivity assessments, and disease infestations (Stone 1998). Since the start of this century, the use of the Global Positioning System (GPS) ground survey and Geographic Information System (GIS) for mapping newly constructed roads and forest harvest boundaries has enabled a powerful spatial analysis to complement aerial photography for forest management in Tasmania. The computer technology and mapping skills have generated highly accurate and sophisticated maps, providing an essential tool for estimating the areas of wood production, and as a result contribute to the management of forests (Elliott et al. 2008).

1.1.3 Remote sensing for forest species distribution and structure mapping

Remote sensing (RS) is a useful set of technologies for assessing, mapping and monitoring landscapes at a range of spatial and temporal scales generating thematic maps for the required information (Gómez et al. 2016). RS technology recording of the reflectance of the Earth's surface can provide useful information about the spatial distribution of forest species (Lee 2008). For example, RS has been used to rapidly depict the loss of biodiversity through land clearing (Skidmore et al. 2015). Remote sensing data can be obtained from a wide range of sources from satellites to drones and used for various environmental applications (Colin et al. 2018). These data have been increasingly used for monitoring forest biodiversity at multiple spatial scales (Ceballos et al. 2015; Lopatin et al. 2016; Zhao et al. 2018). For example, high resolution remote sensing data can be used to generate species distribution maps along with structural information for multiple purposes (Singh et al. 2010).

1.1.3.1 Remote sensing of forest species distribution mapping

Remote sensing technology offers great advantages for acquiring vegetation parameters over large areas without the need to visit sites (Colin et al. 2018; Luo et al. 2016b). Remote sensing technology can potentially be used for identifying the species composition of forests, and species-level identification and discrimination can be conducted to monitor changes in species richness and the succession process of the ecosystem (Sobhan, 2007). This may help forest managers quantify the different tree species available within a forest with higher accuracy in a short time period compared to traditional field inventories. This leads to improvements in environmental sustainability through better management and protection of native vegetation communities (Shang, 2013). This technology offers important opportunities for acquiring species information (Mirik et al. 2013; Shang and Chisholm 2014) in a fast and accurate way, particularly for the analysis of large geographical areas (Dalponte et al. 2008), and to identify hotspots and predict changes in species composition in shortened time and costs (Rocchini et al. 2015). Remote sensing technology, including the rapidly developing multispectral imagery and LiDAR data sources, is broadly applied to identify and quantify forest stand characteristics and allows for better planning, and modelling; for example, it can guide harvesting planning (Alam et al. 2012).

Hyperspectral and LiDAR remote sensing have emerged as powerful tools for ecosystem studies (Higgins et al. 2014), and both types of data can be used to predict species richness in temperate forests (Leutner et al. 2012). Hyperspectral remote sensing, also called imaging spectroscopy, provides fine spectral resolution bands (Chen et al. 2009), through the measurement of the reflected electromagnetic spectrum ranging from visual to the shortwave infrared region (Ghosh et al. 2014). Most importantly, narrow band vegetation indices derived from hyperspectral data reduce atmospheric and water absorption, and the saturation problem contained in broad band vegetation indices typically derived from multispectral satellite data (Chen et al. 2009). Hyperspectral data have been widely used for vegetation species classification that contains detailed information on biophysical and biochemical parameters (Chen et al. 2009; Trier et al. 2018). LiDAR, also called airborne laser scanning (ALS) is an active remote sensing technology, in which return times for laser emissions are used to calculate the elevation of the terrain and the features above ground level (Higgins et al. 2014). In forested ecosystems, laser pulses penetrate through a forest canopy to lower

vegetation layers and the ground, and therefore have great potential for describing canopy structural characteristics (Ahmed et al. 2015; Luo et al. 2016b; Zhao et al. 2018).

Hyperspectral and LiDAR data complement each other, especially when researchers face problems in separate complex vegetation classes (Khodadadzadeh et al. 2015). However, they are still underused within biodiversity research due to discontinuity, unaffordability, and inaccessibility (Skidmore et al. 2015; Turner et al. 2015). The combination of airborne LiDAR and hyperspectral data have yielded promising results indicating the benefits of integration of RS data for species mapping (e.g. Asner et al. 2008; Dalponte et al. 2008, 2012; Féret and Asner 2012; Ghosh et al. 2014; La et al. 2012; Naidoo et al. 2012). LiDAR data can be used to map vegetation structure metrics for habitat quality monitoring and hyperspectral remote sensing has demonstrated its ability to map species composition with high accuracy (Zlinszky et al. 2015).

1.1.3.2 Remote sensing for forest structure mapping

Forest structure denotes the three-dimensional (3D) organisation of individual trees (Masek et al. 2015) that can be simplified into measures of the vertical structure (e.g. density of vegetation layers, canopy heights) and horizontal structure (e.g. canopy cover, leaf area index) (Rutten et al. 2015; Zimble et al. 2003). Forest structure influences wildlife habitat selection, reproduction, and survival (Clawges et al. 2008; Wood et al. 2012) making it an important characteristic for wildlife habitat quality (Cody 1981; MacArthur and MacArthur 1961) including avian biodiversity (Culbert et al. 2013). Timely and verifiable information on forest structure is required for appropriate planning, conservation and sustainable management of forests (Dube and Mutanga 2015; Matasci et al. 2018; Tomppo et al. 2008; Zald et al. 2014), policy formulations, and reporting obligations (Bolton et al. 2015; Zald et al. 2016) aiming to preserve and maintain ecosystem services for ecological, economic and societal needs (Zald et al. 2014; Zald et al. 2016). Measuring forest structure is also vital for biodiversity studies, bushfire modelling, and carbon stock estimation (Hyde et al. 2005). Forest structure is considered a useful biodiversity habitat indicator and has been identified as a key component for monitoring and reporting biodiversity change (Pereira et al. 2013). Mapping of habitat types (Räsänen et al. 2014) and measurement of their structure can contribute to the predictions of species assemblages and their richness, and similarly, its monitoring can help for forest harvesting management as well as landscape level planning (Simonson et al. 2014a).

Remote sensing of forest structure has proven a challenging task for forest managers and planners who largely still depend on aerial photograph survey to meet the users' requirements (Gebreslasie et al. 2010). Forest structural attributes derived from remote sensing include, for example, canopy vertical distribution, stem density and stand height (Masek et al. 2015). Their prediction using multispectral remote sensing is based on empirical relationships between spectral information, such as wavelength and vegetation indices and field measured data (Gebreslasie et al. 2010). Multispectral sensors provide an integrated measurement of structural information on forest height and vertical distribution of foliage. They cannot detect features underneath areas of dense canopy cover and they do not directly provide vertical information on the vegetation attributes. Consequently, they have difficulties distinguishing between canopy layers and ground covers (e.g. grasses vs trees) (Arroyo et al. 2010). They have limited sensitivity to assess vertical as well as understory vegetation structure (Lu 2006). Even hyperspectral remote sensing data has limited ability to capture structural complexity (Zhang et al. 2011). In this context, airborne LiDAR data describes the vertical and horizontal structure of forest vegetation (Haywood and Stone 2011). LiDAR does not have a saturation problem even at high biomass levels (Patenaude et al. 2005), and is not influenced by the presence of cloud, shadows or daylight, and therefore affords more hours for acquisition (Holopainen et al. 2015).

1.1.4 Object-based image analysis for tree crown segmentation and species classification

One mechanism for extracting useful information from remote sensing data is to segment complex imagery into objects. Object-Based Image Analysis (OBIA) in remote sensing and to some degree in GIScience offers many possibilities by providing tools that emulate human perception by grouping homogenous pixels into meaningful image-objects (Chen et al. 2012; Ming et al. 2015). In addition to the original pixel values, OBIA can also incorporate complex derivatives describing the texture, context, and shape (Lopatin et al. 2015), which is helpful for situations where spectral properties are not unique (Blaschke et al. 2014). OBIA provides an effective image processing workflow that derives and classifies real-world objects from remotely sensed imagery. Elevation and intensity data can be appropriately used to create image objects based on user-defined rule sets (Sugumaran and Voss 2007). OBIA can process the spatial and spectral properties of multiple remote sensing datasets with fewer sampling units than other methods (Radoux et al. 2011). Pixel-based classification can

lead to a “salt and pepper” effect in heterogeneous areas, which can be overcome by OBIA decomposing imagery into relatively homogeneous areas and then classifying these areas instead of pixels (Zhang 2014). OBIA consists of two main steps. The first step is image segmentation where the image is decomposed into certain segments; and in the second step, the classification is applied to these segments (Darwish et al. 2003). When image data is acquired at the high spatial resolution, object-based image analysis approaches have proven to be superior to pixel-based approaches (Immitzer et al. 2012).

OBIA represents a trend of processing high-resolution remote sensing imagery (Arvor et al. 2013) and has gained in popularity as an alternative to traditional pixel-based approaches (Mui et al. 2015). Object-based approaches allow for mapping complex, hierarchical habitat systems (Strasser & Lang, 2015). OBIA helps to delineate tree crowns (Hirata et al. 2014; Kuyah et al. 2012) via tree crown segmentation and associated objects that reflect real-world features of interest (Mui et al. 2015). The aim of image classification is to transform continuous data into categorical information classes describing the landscape which can be used for decision making for effective management of natural resources (George et al. 2014). Tree species are the main building block of almost all forest ecosystems, and sustainable management of any forest ecosystem requires a comprehensive understanding of species composition and distribution (Nagendra 2001). The accurate mapping of forest species (e.g. Alonzo et al. 2014; Matsuki et al. 2013; Tagliabue et al. 2016) plays an important role in the sustainable management of forests from an environmental and economic perspective (Matsuki et al. 2015). In Australia, many eucalypt species exhibit similar biophysical characteristics in the context of using remote sensing for the classification of native forest species (Alonzo et al. 2014). For remotely sensed image classification it is therefore preferred to group individual eucalypt species into a broader species group to reduce the chance of spectral mixing and class confusion (Goodwin et al. 2005). Within this context, the use of hyperspectral data and machine learning algorithms has great potential for mapping certain species that lead to the improvements in environmental sustainability using better management of native vegetation communities (Shang and Chisholm 2014).

1.1.5 Modelling forest structural attributes

1.1.5.1 Forest structural attributes

Sustainable forest management requires accurate information on a range of forest stand attributes (Tompalski et al. 2015). Those attributes of forest stands in relation to forest biodiversity, are of special interest (Maltamo et al. 2005). Detailed tree attributes are critical for effective management and analyses of the forest (La et al. 2012) such as modelling ecological heterogeneity to estimate species diversity using remote sensing (Rocchini et al. 2015). Canopy height, canopy cover and vertical canopy structure are considered to be the primary descriptors that characterize the forest structure (Woodgate et al. 2015). Tree heights are often measured in ecological studies characterizing the life history of individual tree species and populations (King and Clark 2011) that are directly associated with the vertical elements of forest structure (McElhinny et al. 2005) and measured by distinct strata (Scanlan et al. 2010). Their quantification is vital for species interactions and biological diversity (Asner et al. 2013). Eucalypt forests comprise a mosaic of structural types (regenerating saplings, regrowth poles, and mature and senescent trees) that reflect a history of timber harvesting and fire (Wood et al. 2017).

Forest inventory is important for a range of forest structure attributes for ecological functions, biodiversity assessments, and tree growth determination (Côté et al. 2009). Remote sensing studies can measure attributes of vegetation structure that vary with the sensors used, and therefore their ability to record and monitor attributes of vegetation (Lawley et al. 2016). Remote sensing technology has become an accepted alternative to mapping various ecosystem properties accurately with standardized processing (Zlinszky et al. 2015), and in the case of satellite remote sensing provide natural resource managers with near-real time data that support conservation efforts (Gillespie et al. 2008). In forest inventory, variation in the densities of reflected laser pulses affects both the bias and accuracy of the predicted forest structural attributes (Magnussen et al. 2010). Over the last two decades, LiDAR technology has emerged as a means to derive a suite of forest structural attributes via mapping and quantifying 3D vegetation structure (MacArthur and MacArthur 1961). LiDAR has been used mainly for the analysis of forest canopy and underlying terrain attributes (Pearse et al. 2019), but is also used for habitat structure and forest biodiversity assessment (Bar-Massada and Wood 2014). Models of habitat quality and spatial patterns of prevalence can be acquired using canopy vertical structure metrics (Simonson et al. 2014a). Armston

(2013) proposed further research on the estimation of canopy structure attributes for forest and woodlands of Queensland or other areas to improve the parameterization of within-crown *Eucalyptus* structure.

Focusing on airborne LiDAR and/or hyperspectral data for tree species classification and forest structural attributes, the relevant peer-reviewed papers are listed in Appendix A-1 and Appendix A-2 published during the last decade.

1.1.5.2 Non-parametric random forest analysis approach

There are numerous ways of analysing remote sensing data, but many involve classification. Recently, machine learning methods have come to dominate the field for developing classifiers from large and complex data typical for remote sensing classification problems. Random forest (RF) is commonly used in this context – it is widely considered to be a relatively simple to use, but highly effective machine learning tool for these purposes. RF is a decision-tree based ensemble classifier (Zhang 2014) which is widely used for both regression and classification problems. Advantages of random forest compared to other statistical classifiers include (i) very high classification accuracy; (ii) a novel method of determining variable importance; (iii) ability to model complex interactions among predictor variables; (iv) flexibility to perform several types of statistical data analysis, including regression, classification, survival analysis, and unsupervised learning; and (v) an algorithm for imputing missing values (Cutler et al. 2007). In previous studies, random forest regression models have been developed to predict stand density, top height, basal area and timber volume (Pearse et al. 2019). The random forest can be used to learn complex non-linear relationships, such as those present in variable vertical forest structure and the association of overstory to understory forest vegetation. The random forest has already been revealed to be very effective for accurate land cover mapping across complex and heterogeneous landscapes (Mellor et al. 2013).

1.2 Problem statement

Tasmanian wet eucalypt forests are internationally important for conservation (Balmer 2016). They also provide the primary sources of native forest timber in Australia and have complex understories (Wood et al. 2015b). The current mapping of forest structure in these systems is based on visual interpretation of stereo images of aerial photographs (Wood et al. 2017). Forest managers and planners used the resulting photo interpretation type maps to describe stand structure of polygons (usually 10's of

hectares in area) relying on photo interpretation skill and known harvesting history of stands (Stone 1998). The results are subjective and generally categorical and are therefore hard to objectively validate and monitor consistently. Photo interpretation type mapping is expensive and time-consuming, leading to a demand for a new approach for identifying species or communities (Stone 1998). Importantly, the workforce of photo interpreters is declining, and PI-type maps are therefore becoming dated. However, the traditional technique of PI-type maps has not yet been fully replaced by the modern remote sensing technologies for mapping forest species distribution and structure.

Wet eucalypt forests contain a dense understory, which is at least partly covered by overstory eucalypts, which makes it difficult to map using aerial photography and other traditional explanatory variables (Tuanmu et al. 2010; Wing et al. 2012). It is also important to understand the interaction of understory structural development and topographic attributes for sustainable forest management practices (Amiri et al. 2017; Tuanmu et al. 2010). Direct forest inventory provides highly detailed information of tree species and forest structure, but that is limited to the sample plots and is costly and time-consuming, particularly for mapping complete spatial coverage of forest structure (Zald et al. 2014) in tall wet eucalypt forests of remote and large geographical areas. Due to tall and dense eucalypt forests, the prediction of understory structure is very poor in southeast Australian forests (Lee and Lucas 2007). The predicted understory properties that are required to manage a wet eucalypt ecosystem are of high societal value because they are rich in biodiversity and are linked to the likelihood and intensity of wildfire (Morsdorf et al. 2010). New research is therefore required to investigate the applicability of LiDAR and high-resolution multi- and hyperspectral imagery and develop robust models to map the presence and structure of understory vegetation (Amiri et al. 2017; Eskelson et al. 2011; Suchar and Crookston 2010).

In addition to hyperspectral and multispectral remote sensing technology, which are mostly used for forest species distribution mapping, LiDAR technology is largely applied for forest structure mapping, but is expensive and mostly limited to a single survey of specific areas. Because a single remote sensing technology cannot provide all the required information, research of low-cost data acquisition for large geographical areas is of high significance to extend the scope of LiDAR-derived variables (McInerney et al. 2010). Recently, land management agencies, forest industries, and public groups have taken keen interest to integrate LiDAR data with multispectral data to retrieve forest structural attributes (Dash et al. 2016; Wilkes et al. 2015). The research

shows the successful application of remote sensing data depends on an appropriate spatial resolution (Kamal et al. 2014). Exploring an optimum resolution remains an unsolved topic of research, especially in wet eucalypt forests and is challenging for ecologists (Zhang 2007). To date, few studies have compared the predicted models of vertical forest structure developed by combining LiDAR-derived topographic attributes with other sources of remote sensing satellite data (Dash et al. 2016; Johansen et al. 2010; McInerney et al. 2010; Mikita et al. 2013). My PhD thesis focuses on the problem of fusing LiDAR data with hyperspectral airborne or multispectral satellite datasets for forest species mapping and characterisation of vertical forest structure. The novelty of this work is to apply different fusion schemes and pixel sizes to experiment and compare them to identify the optimal set of variables and model parameters.

1.3 Research aim and objectives

The overall aim of this thesis is to use machine learning analyses of multi-source remote sensing data to assess and map species and structural composition of wet eucalypt forest. This study focuses on high-resolution airborne LiDAR and hyperspectral datasets for object-based tree crown segmentation and species classification and provides a comparison of model accuracies of vertical forest structure based on extensive derived variables, also from LiDAR and multispectral remote sensing. The analyses were based on a native wet eucalypt forest at the Warra Supersite, Tasmania, Australia. The specific objectives of this research are:

Objective 1

To assess the capability of airborne LiDAR and hyperspectral data for optimal tree crown segmentation and species classification in wet eucalypt forest. This objective addresses the following specific questions:

a) Which combination of datasets can provide higher accuracies in wet eucalypt forests for tree crown delineation and object-based species classification?

b) What are the optimum segmentation parameters when segmenting wet eucalypt crowns using airborne LiDAR and hyperspectral data?

c) Can understory species be identified by fusing airborne LiDAR and hyperspectral datasets in wet eucalypt forests?

d) Can LiDAR contribute to achieving higher accuracy in object-based classification of wet eucalypt species?

Objective 2

To predict the density of three understory layers of a wet eucalypt forest using LiDAR-derived topographic attributes and geology data to represent actual structural composition. This objective addresses the following specific questions:

a) Which attributes of the input variables (LiDAR-derived topographic attributes and geology) can best predict understory density layers?

b) Which are the optimum parameters (*mtry* and *ntree*) for a random forest regression algorithm for the density of three canopy layers?

c) What spatial resolution is most appropriate to model understory density?

d) Which attributes among the input variables are most useful for predicting understory density?

Objective 3

To assess the robustness of multispectral satellite imagery and LiDAR topographic attributes to predict vertical structure across multiple scales in a wet eucalypt forest. To meet this objective, random forest regression modelling with derivatives from three datasets, i.e. airborne LiDAR, WorldView-3 and Landsat-8 (OLI) satellite data were used. This objective addresses the following specific questions:

a) What is the optimum resolution of remote sensing datasets for predicting vertical forest structure?

b) Which combination of remote sensing datasets can yield the best output for modelling vertical forest structure?

c) Can the fusion of derivatives (vegetation indices, texture features, and topographic attributes) improve the model accuracies and produce robust and transferrable models?

d) Do cost-effective low-density discrete return LiDAR and multispectral remote sensing datasets produce models at the desired accuracy to predict vertical forest structure?

1.4 Thesis structure

There are six chapters in this thesis (see Figure 1.1). Chapter 2 describes the study area and details the datasets used in this research. Chapters 3 to 5 of this thesis represent original research studies presented in a form suitable for submission for publication.

These three chapters address the abovementioned core research objectives and questions. Chapter 6 presents synthesis and conclusions about the overall outcomes, limitations, contributions of the research and suggests future directions for mapping species composition and forest structure in wet eucalypt forest.

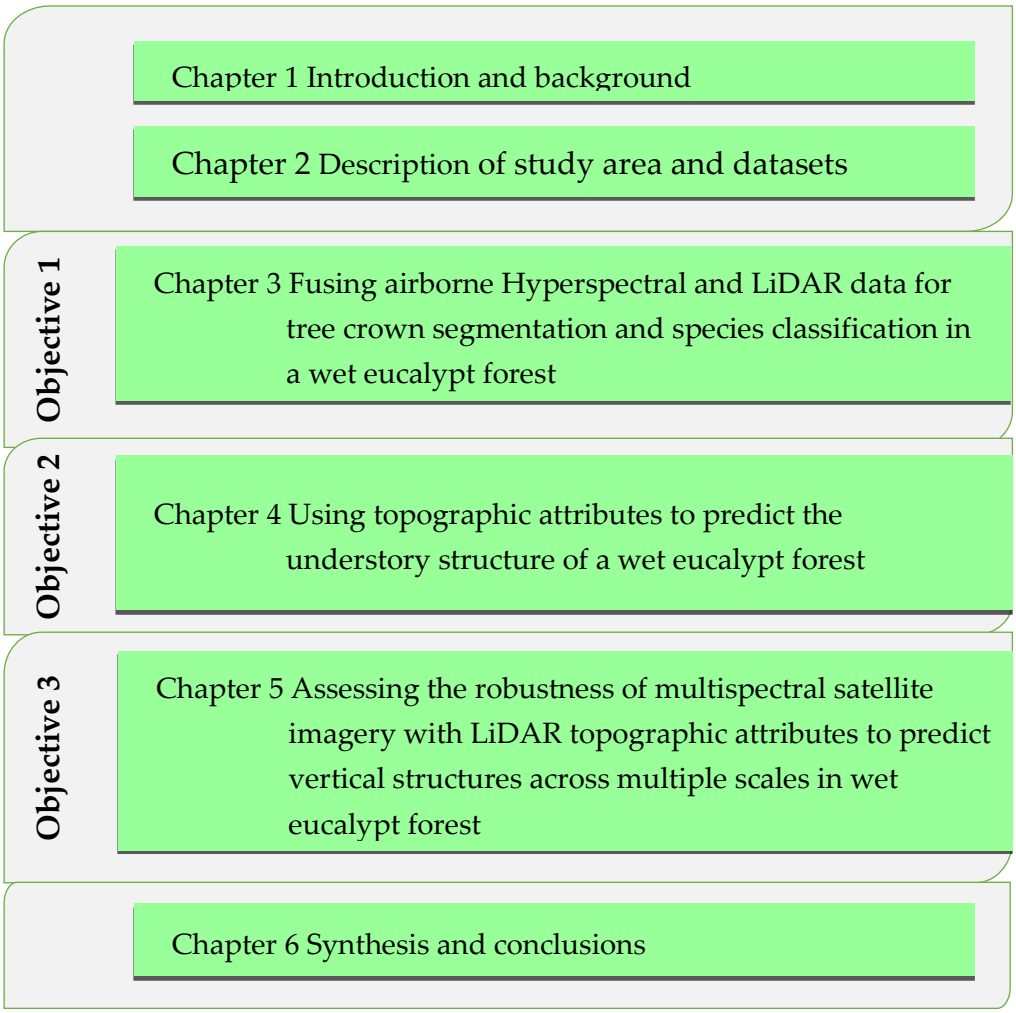


Figure 1.1 Structure of thesis, chapters and research objectives.

Chapter 2

Description of the study area and remote sensing datasets

2.1 Description of the study area: Warra Supersite forests

This study focuses on a 5 km by 5 km area in the region of the Warra Supersite in southern Tasmania, Australia (Figure 2.1). The Warra Supersite is located between the Huon and Weld Rivers, about 60 km west southwest of Hobart. This Supersite has a long history of fire, with several fires since 1850, and ranges from intensively managed to protected forests (Hickey et al. 1999). The Warra supersite is a registered Terrestrial Ecosystem Monitoring Site and is linked to the existing networks of national and international research sites (Brown et al. 2001). There are several permanent research and monitoring plots within the site including a carbon flux tower.

The Warra Long Term Ecological Research (LTER) site of 15,900 hectares was designated in 1995 and established in 1998 to study ecology and monitor one of Australia's productive wet eucalypt forests in Tasmania. The site is located partly in World Heritage Area and partly in Permanent Timber Production Zone which is managed for multiple purposes including timber production. The LTER at Warra has been helping foresters and ecologists to understand ecology and silviculture, and enabling research to contribute to the sustainable management of Tasmanian forests (*https://warra.com/about-us/*). The Warra site is a member of the TERN (Terrestrial Ecosystem Research Network) Australian Supersite[1] Network that represents the cool, temperate wet forest biome. The TERN Supersite network aims to understand the response of key Australian ecosystems to future environmental change and is located in significant Australian biomes that span a wide range of environmental conditions. The Warra Supersite is one of the most productive terrestrial ecosystems in the world, and so its management generates a high level of political and social interest.

The AusCover[2] Programme of TERN (now part of TERN Landscapes) provides important datasets derived from airborne and satellite sensors and focuses on geospatial data products measuring a number of biophysical variables for Australian ecosystems. This programme contributes to the fields of forestry and biodiversity, through a continental scale plot-based monitoring network focusing on forest structure, tree growth, forest productivity and carbon dynamics in tall eucalypt forests across Australia (Johansen and Phinn 2013).

[1] (*https://supersites.tern.org.au/*)

[2] (*http://www.auscover.org.au/about-us/overview/*)

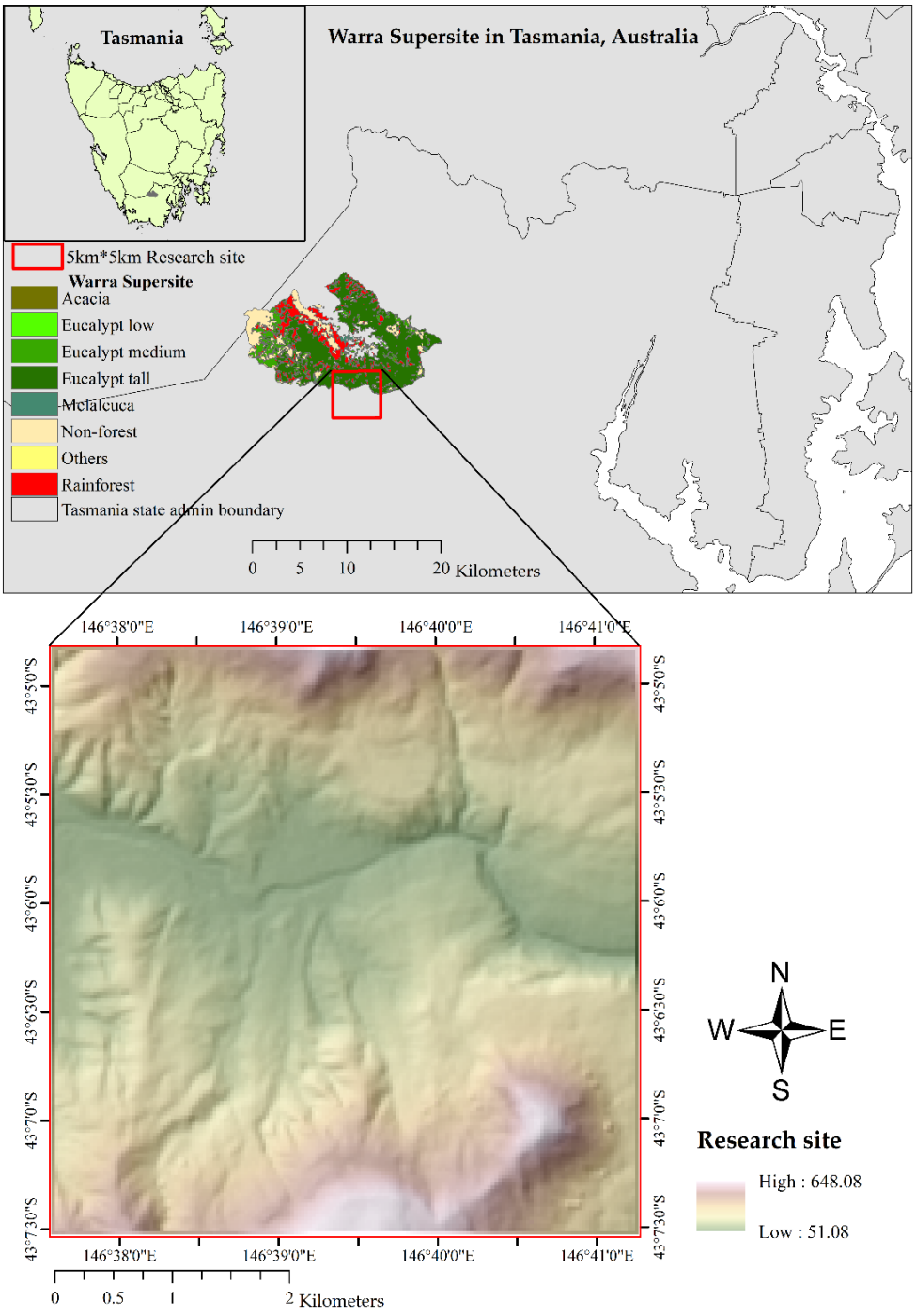


Figure 2.1 Location of the study area. The Warra Supersite is shown in the inset map above and the research site of 5 km by 5 km area. The bottom map shows elevation of the focal 5 km by 5 km study area.

At 495 m above sea level within of the Warra Supersite, mean annual rainfall is 1,707 mm and mean daily temperature ranges are 8.3°C to 19.3°C in January (summer) and 2.5°C to 8.6°C in July (winter) (Bureau of Meteorology 2017). The site, therefore, falls within the temperate climate zone. The ground elevation of the Warra Supersite ranges from 37 m to 1260 m (Brown et al. 2001).

This PhD study was conducted within a 5 km by 5 km forested area located within the Warra Supersite of the Wilderness World Heritage Area in Tasmania, Australia (Figure 2.1). The selected site is topographically complex, comprising harvested sites, rivers, and gullies, and the ground elevation ranges from 51 m to 648 m according to the 30 m DTM.

2.1.1 Species composition and structure

The forest species at Warra are broadly typical of many wet eucalypt forests in Tasmania (Neyland 2001), although species and structural composition differ significantly over small distances (Figure 2.2 – Figure 2.5). The forest at the study site is dominated by *Eucalyptus obliqua*, which forms a tall (~50 m) overstory overtopping an understory containing a range of trees, shrubs and ground-layer species typically classified as wet sclerophyll forest or rainforest. In Tasmanian wet forests, there is a general pattern of ecological succession from wet sclerophyll-dominated younger stands to older mixed forest consisting of eucalypts with a rainforest understory (Baker et al. 2013). Typically for wet eucalypt forests in Tasmania, if the fire interval is <100 years, the understory is likely to be dominated by wet sclerophyll broadleaved trees and shrubs (van Galen et al. 2018). These species are gradually replaced by cool temperate rainforest species where fire interval increases to >100-350 years (Gilbert 1959). The eucalypt species die out if the fire interval exceeds about 350 years, resulting in the rainforest mostly dominated by *Nothofagus cunninghamii* (Hickey et al. 2006).

This study site is suitable for the assessment and implementation of biodiversity conservation strategies. There is a wide range of geologies and soils in southwest Tasmania, with individual species responding strongly to soil nutrition and drainage, and geology being an important driver of plant composition (Neyland 2001). The vertical structure of the wet eucalypt forests contains a mix of species and types, i.e. senescent and mature trees, regrowth poles and regenerated seedlings and saplings that indicate a history of burning and harvesting (Wood et al. 2017). The canopies of tall wet sclerophyll forests are mostly dominated by *Eucalyptus obliqua* over a secondary stratum of trees, for example, *Acacia dealbata* and *Acacia melanoxylon* (Scanlan et al. 2010) or rainforest trees such as *Nothofagus cunninghamii*, *Atherosperma moschatum* and *Eucryphia lucida*. Understory species vary from dense *Melaleuca squarrosa* and *Gahnia grandis* on soils with impeded drainage to *Nematolepis squamea* and *Pomaderis apetala* on well-drained soils (Neyland 2001).

Figure 2.2 Photographs illustrating the interception of solar radiation inside the three layers (Left), and distinction of the upper layer and a lower layer with a middle layer of forest structure (Right).

Figure 2.3 Photographs illustrating the flux tower site with the three layers of forest structure (Left), and fern trees reaching around 10 m in height (Right).

Figure 2.4 A photograph illustrating the old-aged trees with the middle layer of rainforest and the lower layer of fern and other small trees.

Figure 2.5 A photograph illustrating the middle layer of the rainforest with the lower layer of fern and small trees.

2.1.2 Geology and soils

Because of its impact on soils and hydrology, underlying geology plays a crucial role in species and structural composition and growth rates of forests. Geology of the Warra site is dominated by Jurassic and Quaternary slope deposits derived from dolerite

covering ~ 66% of the area (Laffan 2001). These slope deposits overlie Permian sediments across much of the area (Neyland 2001), but Ordovician limestone, Cambrian dolomite, Cambrian volcanic-ultramafic rocks, and Quaternary alluvium and moraine deposits are found in few areas of the site. A geological map (Figure 2.6) of the study area at 1:25,000 scale was made available by Mineral Resources Tasmania[3].

The soil pattern is diverse and shows strong links to geology and vegetation. Where the vegetation cover is forest, mineral soils with gradational texture-profiles are prevalent. Soils formed on quaternary slope deposits derived from dolerite have been described and sampled in some detail from a limited area in the south-eastern part of Warra. There is no high accuracy and precision soil layer for the site. However, it was assumed that the combination of geology and topography should capture a high proportion of important soil characteristics.

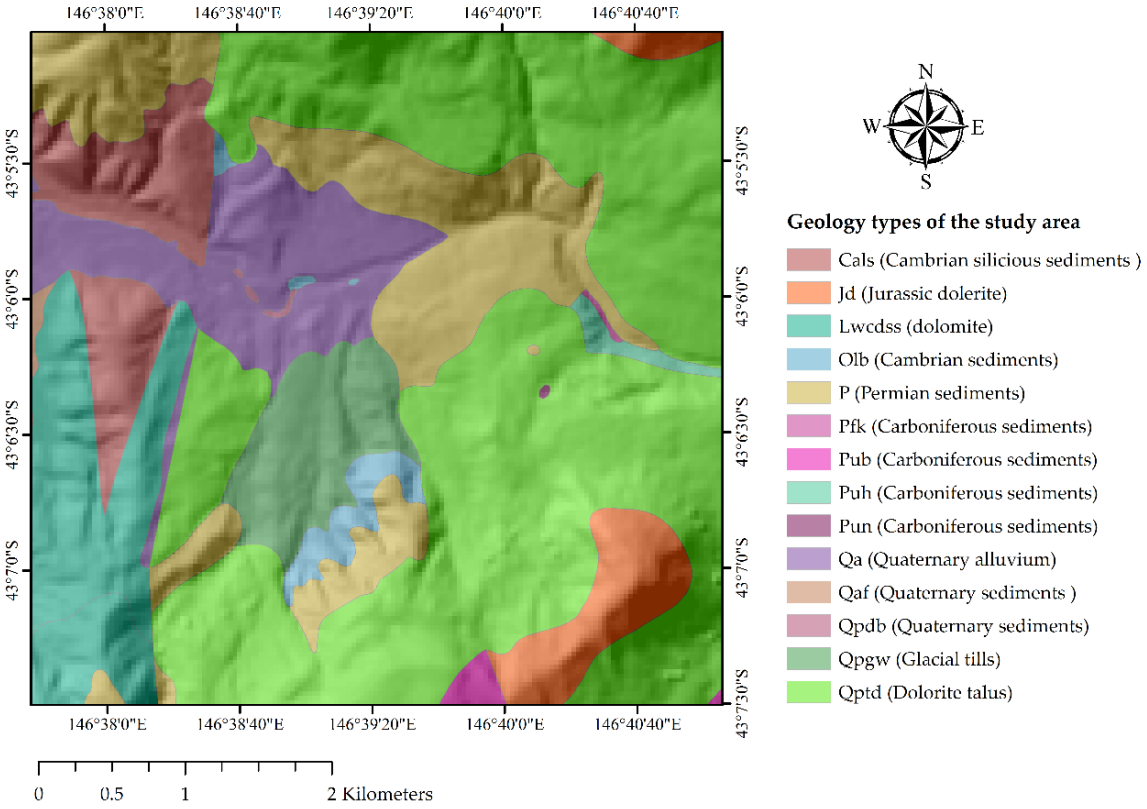


Figure 2.6 Geology map of the study area with hillshade.

2.2 Remote sensing datasets

This thesis is based on an analysis of four types of remote sensing datasets, i.e. airborne LiDAR point clouds, airborne hyperspectral imagery, and satellite WorldView-3 and

[3] *http://www.mrt.tas.gov.au/portal/1-25-000-scale-digital-geology-of-tasmania*

Landsat-8 operational land imager (OLI) multispectral data, to address the three main objectives.

Data for Objective 1: Object-based tree crown segmentation and species classification.

To meet the research objective-1, airborne LiDAR point clouds and hyperspectral imagery (HSI) were used. Technical specifications of the airborne LiDAR and hyperspectral imagery are given in Table 2.1. Field data for training and testing remote sensing datasets for a 1.6 ha permanent vegetation plot were derived from previously made direct field measurements in April 2012 (Wood et al. 2015a). All plants greater than 10 cm diameter at breast height (dbh), tree heights and individual tree coordinates were measured. Since dominant species affect the species composition in the forest ecosystems (Crase et al. 2015), only the four dominant most common plant species, i.e. *Eucalyptus obliqua* (messmate stringybark), *Acacia melanoxylon* (Blackwood), *Dicksonia antarctica* (Soft Tree Fern) and *Nothofagus cunninghamii* (Myrtle beech) were taken into account in this study. The airborne HSI data were collected using an AISA Eagle sensor pushbroom system using a progressive scan CCD with full-width-at-half-maximum (FWHM) ranging between 2.24 and 2.45 nm. The data include blue, green, red and near-infrared bands. Similarly, Airborne LiDAR data used in this study was collected by Airborne Research Australia using a Riegl LMS-Q560 laser sensor on Diamond Aircraft HK36TTC ECO-Dimonas. The scan angle rank values ranged from -36^0 to 44^0. Full-waveform data was collected and discretised up to 7 returns per pulse. The scale factor of the data was 0.01 m. A Microtops II Ozone Monitor-Sunphotometer version 2.46B was used directly to collect information on atmospheric composition during the hyperspectral overflight, which included temperature, pressure, atmospheric water, ozone, and aerosols matching with the nearest flight time of the AISA Eagle sensor. These parameters were applied to the atmospheric correction of HSI data.

Table 2.1 Technical specifications of airborne LiDAR and hyperspectral datasets.

Data	Date of Acquisition	Sensor	Flight height	Spectral range	Spatial resolution	No. of bands
HSI	5-6 February, 2015	AISA Eagle	1387 m	400.71-999.18 nm	0.719 m	252
LiDAR	May 30, 2014	Riegl LMS-Q560	500 m	1064 nm	28.66 points/m^2 (spacing = 0.19 m)	1

Data for Objective 2: Prediction of understory layers using airborne LiDAR-derived topographic attributes.

Airborne LiDAR point clouds and topographic derivatives from the digital terrain model were utilised to meet research Objective-2. A geology data layer was added to investigate the effects of geology for modelling the vertical forest structure.

Data for Objective 3: Assessment of robustness for the prediction of forest vertical structure using airborne LiDAR and multispectral RS data.

To meet research Objective-3, airborne LiDAR point clouds in combination with WorldView-3 and Landsat-8 (operational land imager) multispectral remote sensing datasets were used. This study could not utilise operational scale LiDAR data (e.g. 1 - 8 points/m^2) because such data was not available for the study site for the relevant dates. Downsampling of the high-density research-grade LiDAR data to operational point densities would have included exactly the same points and therefore would not offer a robust comparison as it would artificially inflate predictive power. However, this constraint is unlikely to limit the indirect use of this dataset for analyses derived from DTMs. The high-density LiDAR data were used to derive the response variables of vegetation density in three canopy layers. A simulated operational LiDAR dataset (based on down-sampled point density) was used to simulate a coarser DTM and associated topographic derivatives, which was used to address objective 3.

When the required datasets were compiled for this study, efforts were made to match the acquisition year and month of all the three datasets. WorldView-3 multispectral imagery was provided in TIFF format by DigitalGlobe and was radiometrically corrected. The data provider had orthorectified the multispectral visible, near-infrared (VNIR) and shortwave infrared (SWIR) bands using rational polynomial coefficients (RPCs) with a 5 m LiDAR digital elevation model (DEM) for Z-control, and re-projected from WGS84/UTM55S to GDA94/MGA55. Landsat-8 (OLI) land surface reflectance imagery was downloaded free of cost from the United States Geological Survey (USGS) website. The target path and row were 91 and 90 respectively. The Landsat-8 (OLI) satellite imagery was successfully launched on February 11, 2013, from Vandenburg Air Force Base, California, the USA with a designed life of 5 years. This satellite data offers 16-day repetitive earth coverage and a global archive of sun-lit, substantially cloud-free land images. The satellite carries two sensors, i.e. the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) (USGS 2016).

The technical specifications of multispectral remote sensing datasets are given in Table 2.2.

Table 2.2 Specifications of multispectral remote sensing datasets.

Specification item	WorldView-3 Imagery	Landsat-8 Imagery
Date of acquisition	2015-10-05	2014-10-21
Spatial resolution	1.60 m (VNIR bands) 7.50 m (SWIR bands)	30 m (VNIR and SWIR bands)
Sun azimuth	42.400	48.800454560
Sun elevation	43.700	48.182282610
Product Type Level	"Standard" LV2A	OLI_TIRS_L1TP
Bands (in Nanometres)	Coastal = 427.40 Blue = 481.90 Green = 547.10 Yellow = 604.30 Red = 660.10 Red Edge = 722.70 NIR1= 824.00 NIR2 = 913.60 SWIR1 = 1209.10 SWIR2= 1571.60 SWIR3= 1661.10 SWIR4= 1729.50 SWIR5= 2163.70 SWIR6= 2202.20 SWIR7= 2259.30 SWIR8= 2329.20	Coastal = 442.96 Blue = 482.04 Green = 561.41 Red = 654.59 NIR = 864.67 SWIR 1 = 1608.86 SWIR 2 = 2200.73

Chapter 3

Fusing airborne hyperspectral and LiDAR data for tree crown segmentation and species classification in a wet eucalypt forest

Abstract

To sustainably manage forest biodiversity and monitor changes of species patterning, mapping spatial distribution of tree species is indispensable. Remote sensing (RS) can provide powerful tools for mapping species, but this can be a complex task in areas with high plant diversity, crown closure, and multi-layered canopies. This chapter addresses the issue of classifying temperate wet eucalypt forest by examining tree crown segmentation and species classification using different combinations of RS datasets for a 1.6 hectare plot with mapped tree locations in wet *Eucalyptus obliqua* temperate forest in Tasmania, Australia. This study first explored optimal segmentation parameters to determine the approach with the highest segmentation accuracy compared to the digitized tree crowns. The best segmentation accuracy of 88.71%, resulted from segmenting a combined Minimum Noise Fraction (MNF) dataset derived from hyperspectral imagery (HSI) and the LiDAR-derived Canopy Height Model (CHM). Using a random forest classifier, object-based classification of tree species was then performed for the processed datasets. The fused dataset of MNF and CHM produced the highest accuracy of 78.26% and 66.04% for four and five classes respectively, whereas MNF dataset provided an accuracy of 68.75% and 60% for four and five classes respectively. The fused datasets were found to be more robust to spatially discriminate and classify wet eucalypt forest species compared to a single dataset. This approach classified *Eucalyptus obliqua* with the highest accuracy of 90.86% for four classes using the fused MNF and CHM dataset, and 86.11% for five classes using the fused HSI, indices and CHM dataset. An important understory species - the tree fern (*Dicksonia antarctica*) - was classified with the highest accuracy of 83.54% for four classes using HSI, and that of 84.64% for five classes using the fused HSI and CHM. Therefore, this approach could classify both the overstory and understory species that can play a crucial role in identifying forest biological diversity. The CHM contributed to the accuracy of tree crown segmentation and species classification. Thus, this approach will be useful for forest managers and ecologists for planning sustainable management of wet eucalypt temperate forest biodiversity and producing species maps for monitoring the species of interest. Because "trial-and-error" segmentation approaches are time-consuming, future research should develop the method for automated segmentation of objects.

Keywords: Wet eucalypt forest; Airborne LiDAR; Airborne hyperspectral; Image segmentation; Tree species classification

3.1 Introduction

In recent years, ecologists and forest managers have shown much interest in using remote sensing (RS) data to map individual tree species for biodiversity monitoring (e.g., Baldeck et al. 2015; Bustamante et al. 2015; Dian et al. 2016; Ferreira et al. 2016; Higgins et al. 2014; Kuenzer et al. 2014; Skidmore et al. 2015; Turner et al. 2015; Vaglio Laurin et al. 2014; Yuan et al. 2016; Zhao et al. 2016) and sustainable forest management (e.g., Dalponte et al. 2014; Liu et al. 2011; Wallace et al. 2014; Wulder et al. 2008). Because traditional field inventory assessment of tree species is labour-intensive and time-consuming (Dalponte et al. 2008; Puttonen et al. 2010; Torabzadeh et al. 2014; Xie et al. 2008), RS technology provides a great opportunity for fast and accurate identification and mapping of tree species (Dalponte et al. 2008; Zhang et al. 2016), particularly across large geographical areas and inaccessible sites. Two major types of RS are particularly appropriate for these problems – LiDAR (light detection and ranging) and hyperspectral imagery (HSI).

Hyperspectral imaging provides an advance on earlier spectral imaging approaches available to ecologists for detecting the decline of individual tree species in a particular place (Turner et al. 2003). HSI data contain highly detailed information due to hundreds of narrow contiguous spectral bands from the visible, infrared to the shortwave infrared region of the electromagnetic spectrum. They possess more capacity than multispectral RS data to classify different species (Clark et al. 2005; Dalponte et al. 2012; Dalponte et al. 2014; Ferreira et al. 2016; Liu and Bo 2015). They can sometimes even resolve subtle spectral differences between species (Cho et al. 2012; Ferreira et al. 2016). Thus, HSI data have become a cost-effective tool for areas with high plant diversity (Vaglio Laurin et al. 2014). However, HSI data may fail to detect differences among species of similar spectral signatures (Dalponte et al. 2008; Ghamisi et al. 2014; Jones et al. 2010; Liu et al. 2011) in highly mixed vegetation sites (Dian et al. 2016). Also, HSI mainly captures information on canopy surface features and therefore does not provide vertical structural information (Dalponte et al. 2008; Dian et al. 2016). However, LiDAR provides valuable insight into a tree and forest canopies (Marrs and Ni-Meister 2019) and detailed information about vertical structure but no information about spectral composition (Debes et al. 2014). Thus, LiDAR data complement the information provided by HSI (Dalponte et al. 2012; Dian et al. 2016; Khodadadzadeh et al. 2015; Marrs and Ni-Meister 2019; Zhou and Qiu 2015), and fused HSI and LiDAR data has the potential to provide greater classification accuracy than either HSI or LiDAR data alone (Marrs and Ni-Meister 2019). However, the greater the

number of species, the lesser the classification accuracies in biodiverse complex forests (Dian et al. 2016). In this context, this chapter focuses on the predominant wet eucalypt forest species only.

The object-based image analysis (OBIA) approach offers many opportunities for using spatial and spectral properties of images for classification. Many studies have demonstrated that this approach is superior to the pixel-based classification (PBC) approach (e.g., Immitzer et al. 2012; Kamal et al. 2015; Ke et al. 2010; Zhou and Qiu 2015). OBIA involves grouping homogenous pixels into meaningful image objects (Chen et al. 2012; Ming et al. 2015) based on multiple attributes such as shape, texture, and context (Ke et al. 2010; Lopatin et al. 2015). This approach first segments an image into homogeneous objects and then classifies them into classes (Arroyo et al. 2010; Whiteside et al. 2011; Zhang 2014) that reflect the real-world objects of interest (Mui et al. 2015). Different kinds of image segmentation techniques, broadly termed as bottom-up and top-down segmentation have been introduced for remote sensing applications (eCognition 2014). A bottom-up approach, i.e. multiresolution segmentation (MRS) approach, also called the Fractal Net Evolution Approach (FNEA), is a widely used approach to generate image objects (e.g., Blaschke 2010; Bonnet et al. 2015; Levick et al. 2015). This MRS approach minimizes heterogeneity (starting with individual pixels) and maximizes their respective homogeneity based on a user-defined scale parameter, shape and compactness (Liu and Bo 2015). Nevertheless, exploring optimal segmentation parameters is still an issue for research that combines HSI and LiDAR data for classifying forest tree diversity. Furthermore, the application of OBIA requires the use of classification algorithms. One particularly useful approach is random forest (RF), which is an ensemble non-parametric machine learning classifier originally developed by Breiman (2001). This classifier is robust and can handle high dimensional input datasets without overfitting. Owing to its high capability to classify and characterize complex interactions of variables, this classifier has been widely used by ecologists (Cutler et al. 2007) and is faster than other ensemble classifiers such as support vector machine (SVM) and AdaBoost (Belgiu and Drăguţ 2016).

The wet eucalypt temperate forests of Tasmania are both economically important and important for biodiversity (Brown et al. 2001). Forest planners use GIS layers describing stand structure relying on aerial photograph interpretation (PI) type skill, and known harvesting histories of stands (Stone 1998). Photogrammetrists used stereo images at a scale of approximately 1:20,000 to delineate polygons based on visual interpretation (Wood et al. 2017). However, PI type mapping is expensive and time-

consuming, which leads to a demand for a new approach for identifying species or communities (Stone 1998). This chapter, therefore, assesses the capability of airborne LiDAR and hyperspectral data for optimal tree crown segmentation and species classification in a wet eucalypt forest.

3.2 Methods

In this chapter, tree crowns were segmented and tree species classified with the random forest classifier based on objects, as summarized in Figure 3.1. The segmentation and classification accuracies were then evaluated.

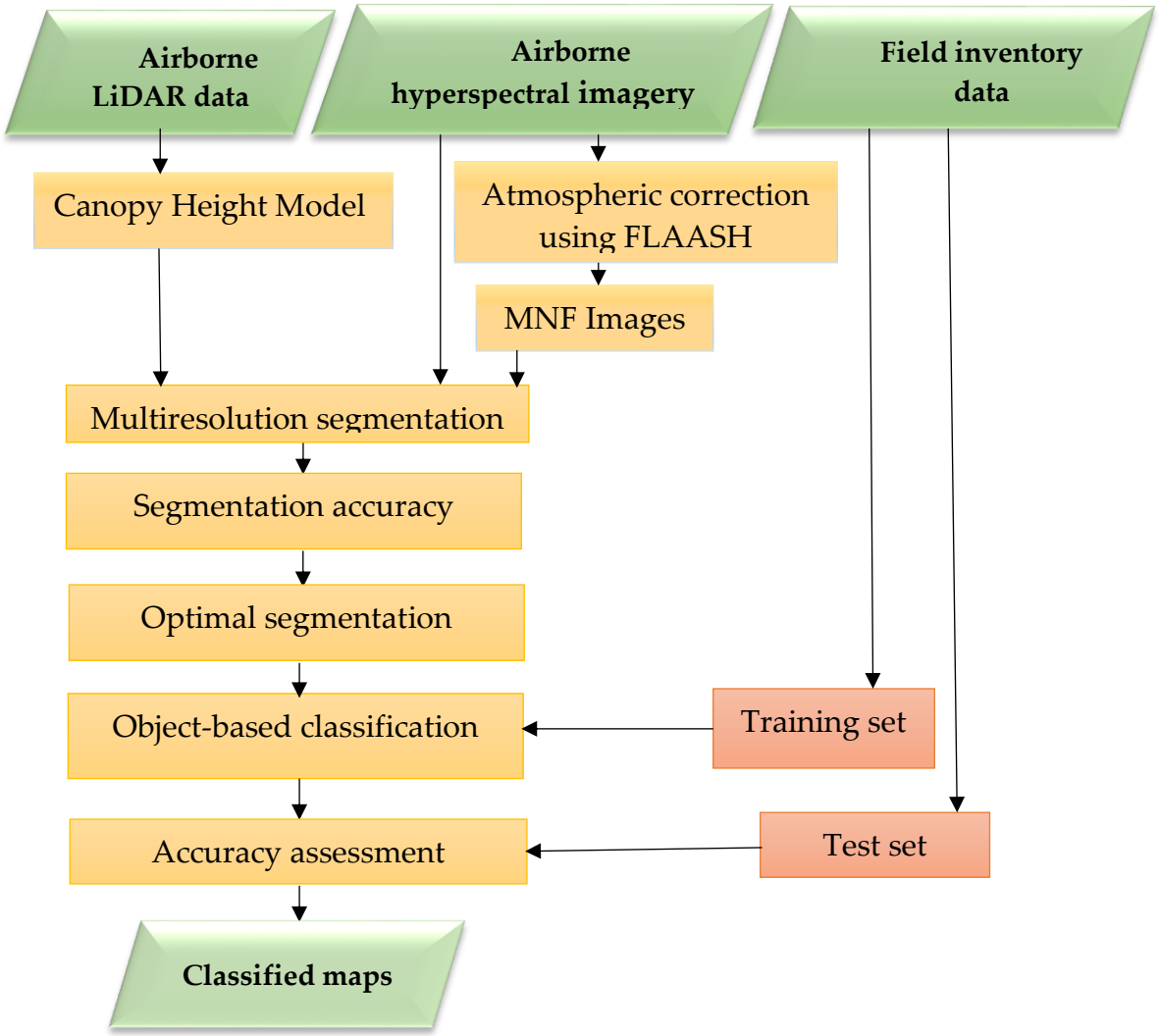


Figure 3.1 Workflow diagram for forest tree crown segmentation using multiresolution segmentation and species classification using the random forest classifier. Airborne LiDAR and hyperspectral data were separately processed for canopy height model (CHM = DSM – DTM) and minimum noise fraction (MNF) images respectively. Training and test datasets were extracted from field inventory data of tree locations.

The training set was applied to train the RF classifier and the test set for cross-validation.

3.2.1 Test site and datasets

This study was carried out in a 1.6 ha research plot situated within the Warra Supersite - a Long Term Ecological Research (LTER) site (Figure 3.2). Airborne LiDAR and HSI data were used. The detail description of the study area and datasets have been given in Chapter 2.

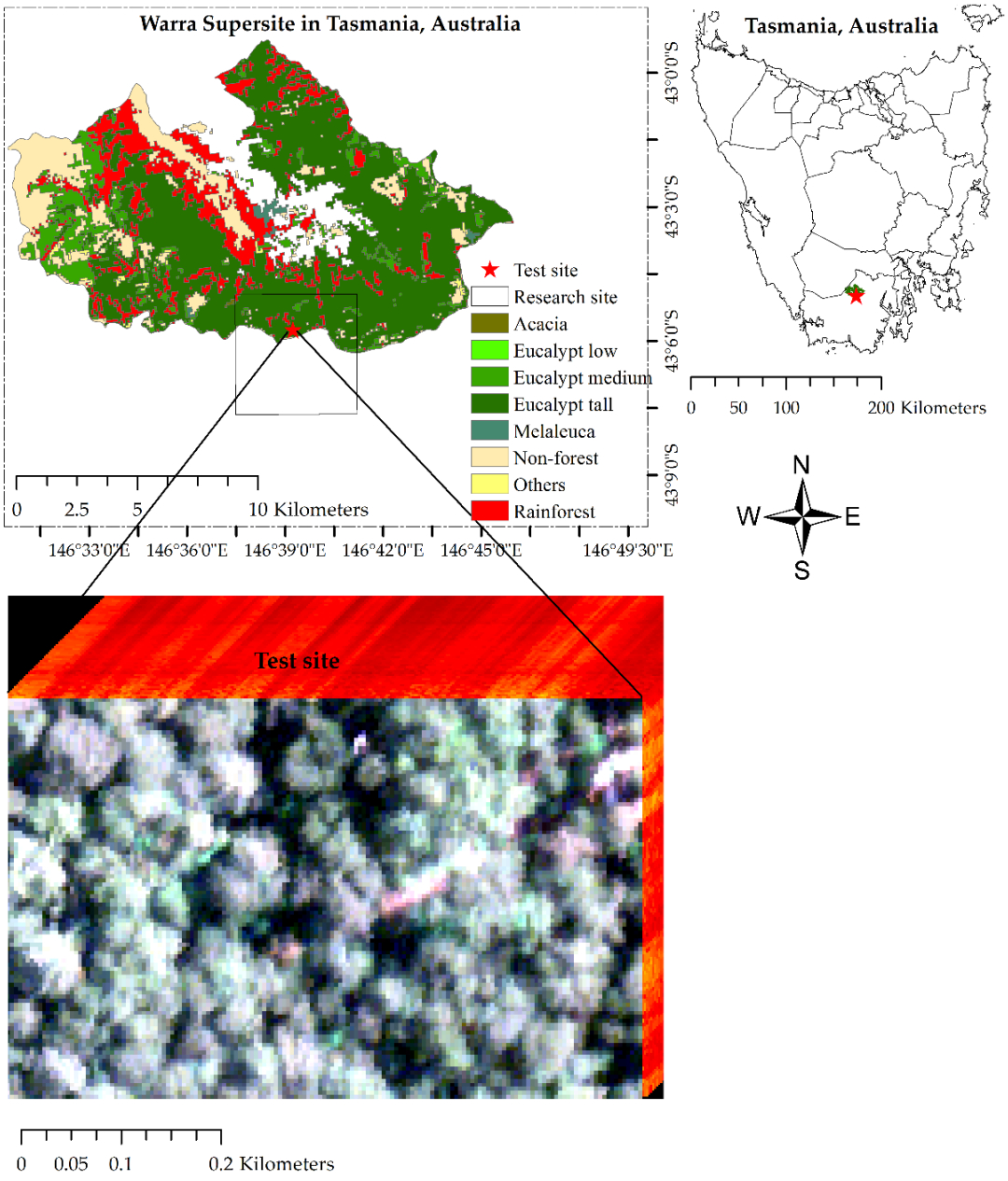


Figure 3.2 Location of study site: the inset map shows the test site inside the Warra Supersite and Warra Supersite inside Tasmania, Australia. The hypercube is derived from airborne hyperspectral data for the test site.

3.2.2 Data pre-processing

The HSI and LiDAR data were processed separately. First, the LiDAR point clouds of the test site were clipped and processed using LAStools software (version 170818). The clipped LiDAR point clouds were classified into the ground and non-ground points. A digital terrain model (DTM), a digital surface model (DSM) and then a pit-free canopy height model (CHM) were derived with the step size of 0.719 m from the LiDAR data using the scripts of *lasground*, *lasheight*, *las2dem*, *lasthin*, and *lasgrid* following Khosravipour et al. (2014). The step size of 0.719 m was chosen to match the spatial resolution of the HSI data.

In the case of HSI data, the test site was clipped from the original radiance data. HSI data can be impacted by noise and high dimensionality. Atmospheric effects caused by aerosol and molecular scattering were first minimized to retrieve an approximation of the accurate surface reflectance values (e.g., Brovkina et al. 2016; Dalponte et al. 2012; Dian et al. 2016; Kamal et al. 2015; Zhang et al. 2013b). The Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) was applied using the atmospheric model input parameters, i.e. atmospheric model, sensor altitude, aerosol model, initial visibility, water column multiplier, aerosol retrieval, spectral polishing, and modtran resolution, in ENVI software version 5.3.1. Then, an MNF transformation was performed to denoise the spectra and reduce the dimensionality of the atmospherically corrected reflectance image dataset (e.g., Ghosh et al. 2014; Voss and Sugumaran 2008; Zhang 2015; Zhang et al. 2016). Because MNF transformation first decorrelates and re-scales the noise, and then uses principal components derived from the original data, this transform is considered superior to other methods, such as Principal Component Analysis (PCA) (Batini et al. 2017; Gao et al. 2017; Luo et al. 2016a; Marrs and Ni-Meister 2019; Nielsen 2011). The first four MNF bands were selected for tree crown segmentation that was visually free from noise (Green et al. 1988). The four representative bands were selected from HSI data, i.e. the blue band (B; centered at 446 nm), green band (G; centered at 533 nm), red band (R; centered at 659 nm) and near-infrared band (NIR; centered at 872 nm) to compare with four MNF bands to examine how they perform for tree crown segmentation and species classification (Liu and Bo 2015). For that, all the spectral bands of hyperspectral imagery were grouped into four clusters (R, G, B, and NIR), and the representative centered four bands were selected from these clusters (Yang et al. 2017a).

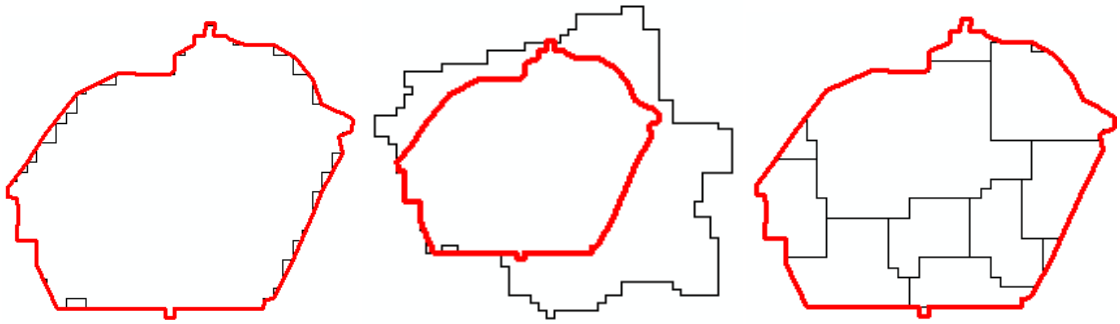
3.2.3 Tree crown segmentation approach

This study deployed a widely used MRS approach (e.g. Blaschke 2010; Bonnet et al. 2015; Levick et al. 2015; Liu and Bo 2015) to segment tree crowns based on the four combinations of datasets, i.e. HSI, HSI/CHM, MNF, and MNF/CHM. The MRS approach aims to delineate individual tree crowns for species classification (Ballanti et al. 2016). This approach requires three parameters, i.e. a scale parameter (to modify the size of objects), shape (to control the influence of object shape and spectral information on segment formation), and compactness (to determine the shape of the segments, which is influenced by the compactness and smoothness parameters) (Ballanti et al. 2016; Liu and Bo 2015). Moreover, the scale parameter controls the dimension and size of segmented objects. The optimizing scale has become an important issue for research in OBIA (Ma et al. 2017). Shape (textural homogeneity of the resulting image objects) and color (digital value of the resulting image objects) parameters range from 0.01 to 1.0 (eCognition 2014). The weight for the CHM layer was set to 30 to emphasize the importance of tree height, and a weight of 1.0 was used for all other layers (Liu and Bo 2015). Previous studies have shown that the better the matching of segments with reference data, the greater the classification accuracies could be obtained (Dian et al. 2016; Liu and Bo 2015). Individual tree crown segments enable us to derive spectral data for each crown to assist with species classification (Pouliot et al. 2002). The geometric accuracy of the segmentation needs to be quantified to optimize segmentation parameters. Accurate segments match more than 50% with reference objects (Clinton et al. 2010), whereas over-segmentation indicates that a reference object is segmented into two or more segments, and under-segmentation indicates that the segmentation has resulted in segments that are bigger than the reference objects (Figure 3.3). Segmentation accuracies were assessed using the visually selected 100 segments at each level of scale.

This chapter first focused on achieving optimal segmentation parameters performing the MRS approach (Table 3.1). Scale parameters ranged from 5 to 45 that were iteratively tested. Segments below a scale of 5 were too small to assess segmentation accuracies, and above a scale of 45, there were no more than two segments for the whole image, which was too coarse.

Table 3.1 Segmentation schemes applied in OBIA for determining optimal segmentation parameters.

Segmentation scheme range			Increment			Total number
Scale	Shape	compactness	Scale	Shape	compactness	
5 - 9	0.05 - 0.95	0.05 - 0.95	1	0.1	0.1	5×10×10
10 - 19	0.05 - 1.0	0.05 - 1.0	1	0.05	0.05	10×20×20
20 - 45	0.05 - 0.95	0.05 - 0.95	5	0.1	0.1	5×10×10



(**a**) Accurate segmentation (**b**) Over-segmentation (**c**) Under- segmentation

Figure 3.3 Examples of accurate, over- and under-segmentation (a-c). Reference polygons are shown in red, and segmented objects are shown in black.

3.2.4 Reference data collection

Because pit-free CHM derived from LiDAR data is useful for overlapping tree crowns in dense forests (Leckie et al. 2003), all the tree crowns of the 1.6 ha test site were delineated manually using the pit-free CHM to use as reference segments for segmentation accuracy. Stratified random sampling was used to choose training and test sets for tree species classification and accuracy assessment (e.g., Ballanti et al. 2016; Duro et al. 2012). Approximately two-thirds of tree locations were selected for training and one-third for validation (Breiman 2001; Duro et al. 2012) (Table 3.2 and Figure 3.4). The training sets were applied to train the random forest classifier, and test sets to validate the object-based classification.

Table 3.2 Training and test sets for classification and accuracy evaluation.

Class name	Total number	Randomly selected tree locations for	
		Training set	Test set
Eucalyptus obliqua	223	97	45
Acacia melanoxylon	156	28	15
Dicksonia antarctica	540	29	16
Nothofagus cunninghamii	181	23	11
Bare ground and non-vegetated objects	78	11	5
Total	1100	188	92

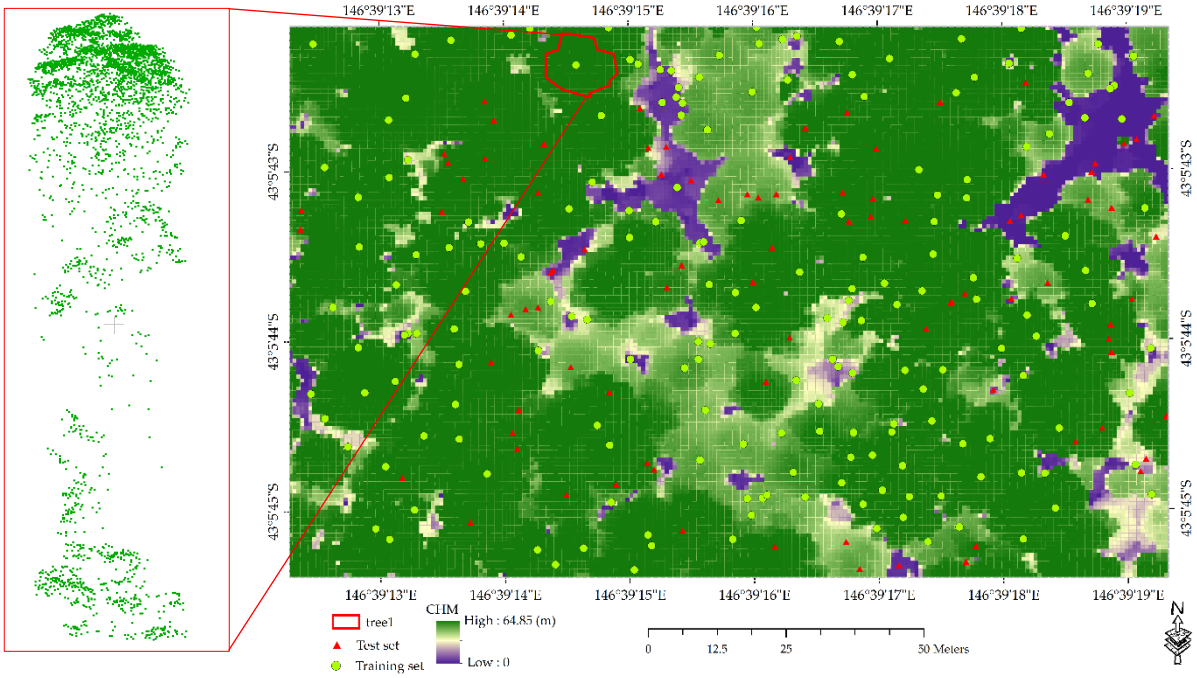


Figure 3.4 Test site showing training (light green) and test sets (red) on CHM (right) and LiDAR point clouds of a tree (left).

3.2.5 Classification approach

3.2.5.1 Object-based classification using random forest

The image objects (corresponding to tree crowns or major parts of crowns) were classified based on the random forest classification algorithm (Dalponte et al. 2012; Ghosh et al. 2014; Merentitis et al. 2014; Naidoo et al. 2012; Onojeghuo and Onojeghuo 2017; Ruiz et al. 2016; Zhang et al. 2016). This machine learning classifier is a supervised technique (Di Lallo et al. 2017) that starts with the selection of many bootstrap observations with replacement (meaning the same samples may be selected

several times) from the dataset. Approximately 63.2% of the datasets are used as bootstrap samples, and the rest of them (36.8%) are used as the test dataset in a typical dataset (Cutler et al. 2007). The training and test tree locations were converted into objects using eCognition developer software and exported to shape files using R programming language for object-based classification. This study used the following six combinations of datasets for object-based classification.

1. Hyperspectral image bands (HSI)
2. Hyperspectral image bands + CHM (HSI/CHM)
3. Reduced MNF hyperspectral bands (MNF)
4. MNF + CHM (MNF/CHM)
5. HSI + vegetation indices (HSI/Indices)
6. HSI + indices + CHM (HSI/Indices/CHM)

The classification scheme also tested four vegetation indices (VIs), adding them to the HSI dataset to improve classification accuracy. The Normalized Difference Vegetation Index (NDVI) (Tucker et al. 1984; Tucker 1979), Simple Ratio Vegetation Index (SRVI) (Jackson and Huete 1991), and Green Ratio Vegetation Index (GRVI) serve to highlight the presence of dense and chlorophyll-rich vegetation canopies. The Leaf Area Index (LAI) (Viña et al. 2011) is a strong indicator of biomass and vegetation cover. The selected VIs were calculated using Equation (1) (Jackson and Huete 1991), Equation (2) (Tewari et al. 2003), Equation (3) (Sripada et al. 2006), and Equation (4) (Jackson and Huete 1991):

$$\mathrm{NDVI} = \frac{\text{Near Infrared} - \text{Red}}{\text{Near Infrared} + \text{Red}} \quad (1)$$

$$\mathrm{LAI} = 0.57 * \exp(2.33 * \mathrm{NDVI}) \quad (2)$$

$$\mathrm{GRVI} = \frac{\text{Near Infrared}}{\text{Green}} \quad (3)$$

$$\mathrm{SRVI} = \frac{\text{Near Infrared}}{\text{Red}} \quad (4)$$

The first six bands of MNF data were selected because these bands contained most of the variance in high-dimensionality hyperspectral data (Liu and Bo 2015; Priyadarshini et al. 2019). The MNF technique is widely used to reduce spectral noise in hyperspectral imagery (e.g. Ghosh et al. 2014; Gomez-Chova et al. 2003; Jones et al. 2010; Zhang et al. 2013a). The dominant species, i.e. *Eucalyptus obliqua, Acacia melanoxylon*, *Nothofagus cunninghamii*, and *Dicksonia antarctica* were selected to evaluate

the classification accuracies. A class "Bare ground and non-vegetated objects" was also added to avoid the holes in the maps covering the test site thus reduced the extra work of masking in this chapter.

Different packages in the R statistical language (R Core Team 2017) were used for classification: '*randomForest*' package for classification based on a forest of trees using random inputs (Liaw and Wiener 2018), '*caret*' package for conversion functions for the class confusion matrix (Kuhn 2017), *'raster'* for analyzing and modeling gridded spatial data, and '*sp'* for constructing the spatial data frame using geometry and attributes (Pebesma et al. 2018), the package '*readr*' is also used to read tabular data (Wickham et al. 2017), and '*ggplot2*' to create graphics for data analysis (Wickham 2017).

3.2.5.2 Random forest parameters and variable importance

Random forest requires only two parameters: "*ntree*" (the number of decision trees grown) and "*mtry* " (the number of variables randomly sampled as candidates at each tree node split) (Breiman 2001; Dalponte et al. 2012; Fassnacht et al. 2014b). The *mtry* is typically set as the square root of the number of predictor variables, and the default value of *ntree* is 500. The *ntree* parameter is adjusted over several iterations of the model to assess how overall classification accuracy is affected by the number of decision trees grown. (Mellor et al. 2012). The trees are fully grown and each is used to predict out-of-bag observations. The number of predictors used to find the best split at each node is a randomly chosen subset of the total number of predictors (Prasad et al. 2006). The predicted class of observation is calculated by a majority vote of the out-of-bag predictions for that observation, with ties split randomly (Cutler et al. 2007).

This study used the total number of variables for *mtry* and the default value of *ntree* based on the studies carried out by Abdel-Rahman et al. (2013); (Ghosh et al. 2014; Reese et al. 2014) and Di Lallo et al. (2017) although the study on optimum parameters requires more detailed consideration, and is therefore considered in Chapter 4.

The variable importance (VI) considers the misclassification rate for the out of bag (OOB) samples for each tree in the forest. Prasad et al. (2006) reported: "variable importance is evaluated based on how much worse the prediction would be if the data for that predictor were permuted randomly". To evaluate the importance of the predictor variable, its values are first permuted randomly from the out of bag samples and then modified out of bag datasets are passed down to the tree to get the new predictions. The difference between those two values divided by the standard error provides the importance of the variable (Cutler et al. 2007; Liaw and Wiener 2018). The

variable importance of a particular predictor variable is calculated using Equation (5) (Genuer et al. 2015).

$$VI(X^j) = \frac{1}{ntree}\sum_t \left(\widetilde{errOOB_t^j} - errOOB_t\right) \quad (5)$$

Where the variable importance of a particular predictor variable is ($VI(X^j)$), each tree of the forest (t), the error of a single tree t on OOB_t (data not included in the bootstrap sample to construct t) sample ($errOOB_t$) (mean square error for regression and misclassification rate for classification), a perturbed sample of OOB_t ($\widetilde{errOOB_t^j}$), and a number of trees ($ntree$).

The variable importance is selected based on Mean Decrease Accuracy (MDA) or Mean Decrease in Gini (MDG). The Mean Decrease in Gini measures Gini impurity for a specific class, whereas Mean Decrease Accuracy evaluates the difference between out of bag error of a dataset obtained from the permutations of the values of variables and out of bag error from the original dataset (Breiman 2001). With the random forest classifier, each and every tree votes for a class, and the classes are finally selected with the majority votes to predict a class (Belgiu and Drăguţ 2016; Di Lallo et al. 2017). Finally, confusion matrices and maps were generated to interpret the results.

3.2.6 Accuracy assessment

To assess segmentation accuracy, the reference objects were used based on the overlapping regions and the difference in areas between the reference objects and the segmented objects they intersect (Clinton et al. 2010). The quality of segmentation was determined using the goodness of fit (D value) measuring over- and under-segmentation. The D value ranges from 0 to 1 (Clinton et al. 2010); the lower the value of D, the greater the segmentation accuracy. The over- and under-segmentation accuracies were assessed using Equations 5 and 6.

$$\text{Over segmentation}_{\text{ij}} = 1 - \frac{area\,(x_i \cap y_j)}{area(x_i)},\ y_j \in Y_i^* \quad (6)$$

$$\text{Under segmentation}_{\text{ij}} = 1 - \frac{area\,(x_i \cap y_j)}{area(x_i)},\ y_j \in Y_i^*. \quad (7)$$

Where,

x_i (i $= 1, 2, \ldots\ldots n$) is the set of training objects or reference polygons, relative to which the segmentation will be judged, and y_j ($j = 1, 2, \ldots\ldots\ldots n$) is the set of all segments in the segmentation. Briefly, $\frac{area\,(x_i \cap y_j)}{area(x_i)}$ are the areas of the geographic intersection of training object, x_i and image segments, y_j. Area (.) is the geographic area of '.', and Y_i^*

is the subset of segments (Clinton et al. 2010). If over- and under-segmentation is found to be zero that means segmented objects match with training objects perfectly. The goodness of fit (D value) was calculated using Equation (8) (Clinton et al. 2010) to select an optimal segment and its parameters

$$D_{ij} = \sqrt{\frac{Over\ segmentation_{ij}^2 + under\ segmentation_{ij}^2}{2}} \quad (8)$$

Approximately one-third of the samples are used to assess the classification accuracies. The random forest classifier uses out of bag error predictions (data not used for training purposes) and then averages all the observations. In addition to overall accuracy (OA) and the Kappa statistic (*k*), this study interpreted accuracies based on sensitivity, specificity and balance accuracy instead of producer's and user's accuracies. Because the conventional point estimate overestimates accuracy because of taking advantage of imbalance test sets, the balance accuracy overcomes the problems of overestimation (Brodersen et al. 2010). This study presented and interpreted the overall accuracy, sensitivity, specificity, and balance accuracy. Sensitivity (true positives) measures the percentage of presences classified correctly, and specificity (true negatives) measures the percentage of absences classified correctly (Cutler et al. 2007; Dehzangi et al. 2015; Xia et al. 2013). Balance accuracy is the average accuracy obtained from sensitivity and specificity (Brodersen et al. 2010). Sensitivity, specificity and balance accuracy are calculated as,

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP})$$

$$\text{Balance accuray} = \frac{1}{2}\left(\frac{\text{TP}}{\text{P}} + \frac{\text{TN}}{\text{N}}\right)$$

Where, TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

3.3 Results

3.3.1 Tree crown segmentation and accuracy assessment

The first goal of this study was to determine optimal segmentation parameters using four datasets. The fused MNF and CHM dataset had a substantial improvement over the other three datasets with an overall segmentation accuracy of 88.7%. The other three datasets produced similar accuracies to each other (48.4% - 54.8%). The results demonstrated an impact of scale, shape and compactness parameters on the accuracy

of tree crown segmentation (Table 3.3) and the D-values (goodness of fit) from the scale of 5 to 45 are presented in Figure 3.5.

Table 3.3 Optimum segmentation parameters using hyperspectral and minimum noise fraction datasets with/without canopy height model, and their D-values (goodness of fit values) and segmented objects' accuracies.

Segmentation scheme	Optimum segmentation parameters			Goodness of fit (D-value)			Segments' Accuracy
	Scale	Shape	Compactness	Over-segment	Under-segment	D-value	
MNF/CHM	10	0.3	0.2	0.0652	0.1256	0.1129	88.71%
MNF	8	0.5	0.2	0.3528	0.4831	0.4522	54.78%
HSI/CHM	25	0.8	1.0	0.3603	0.4911	0.4634	53.66%
HSI	30	0.9	0.9	0.5524	0.3956	0.5160	48.40%

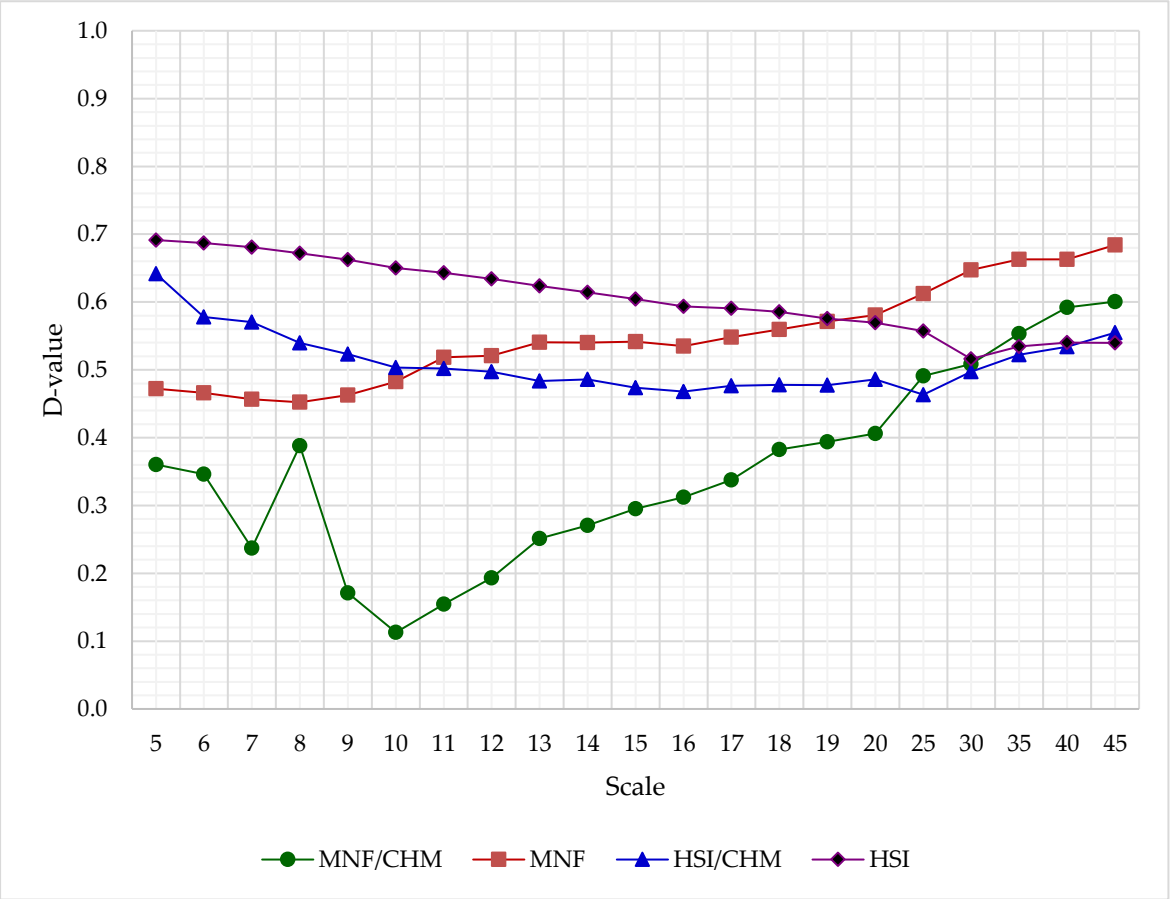


Figure 3.5 Segmentation quality assessment based on D-values using minimum noise fraction and hyperspectral data with/without canopy height model for a range of 5 - 45 of the scale parameter. The lower the D-value, the higher the accuracy.

3.3.2 Object-based classification accuracy

3.3.2.1 Predicted tree species

Based on object-based tree species classification using random forest classifier, crown projection area (CPA- Area projected vertically by tree canopy), mean heights, maximum heights, and minimum heights were predicted. The predicted ranges of the crown projection area, mean heights, maximum heights and minimum heights of individual trees clearly indicate that the forest is very complex with mixed canopies of different species (Figure 3.6).

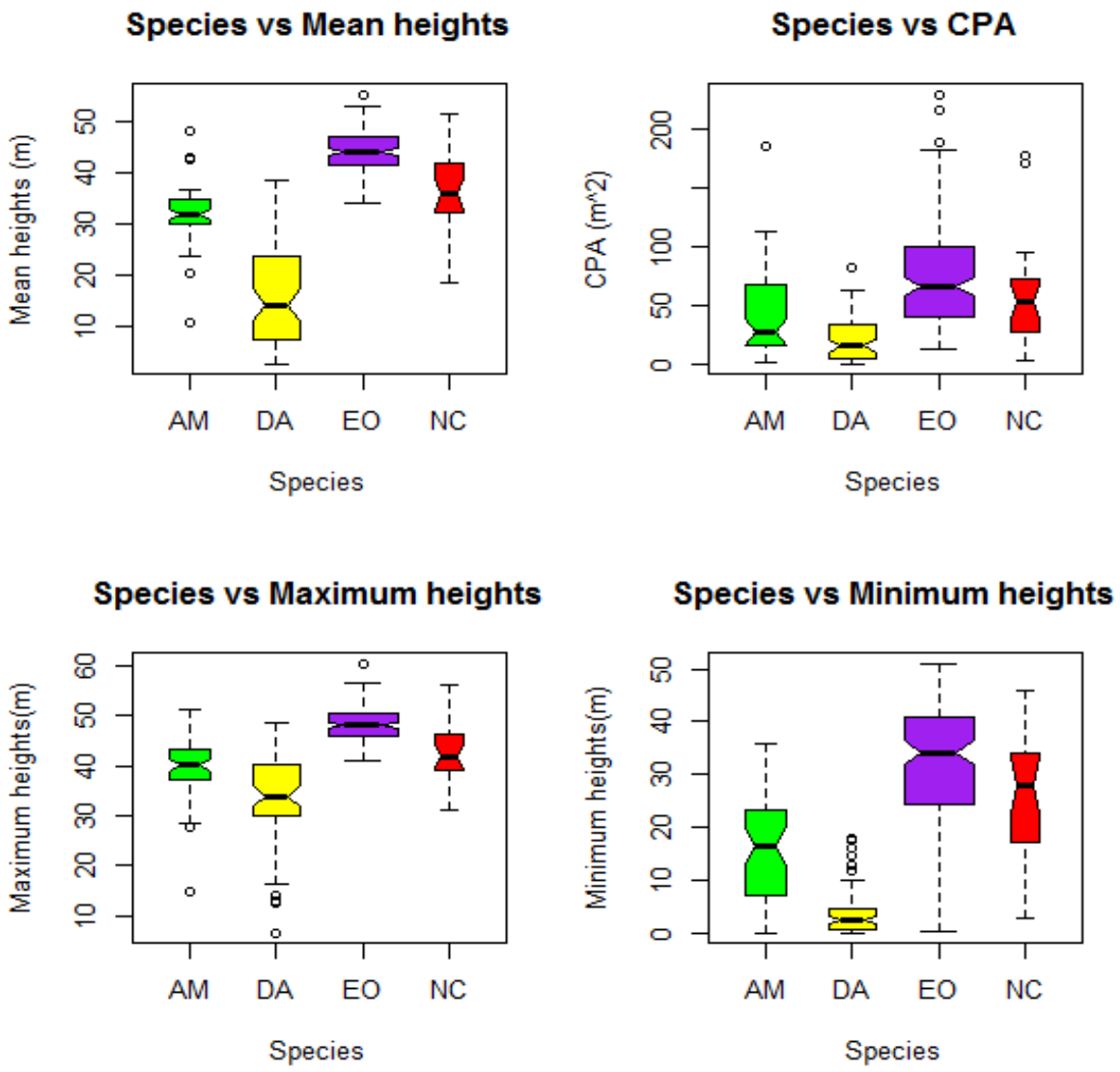


Figure 3.6 Box plots of modelled mean heights, canopy projection area (CPA), maximum and minimum heights of *Acacia melanoxylon* (AM), *Dicksonia antarctica* (DA) *Eucalyptus obliqua* (EO), and *Nothofagus cunninghamii* (NC) are shown based on MNF/CHM dataset.

Among the five prediction classes (i.e. *Eucalyptus obliqua, Acacia melanoxylon, Dicksonia antarctica, Nothofagus cunninghamii* and Bare ground and non-vegetated objects), *Eucalyptus obliqua* as an object possessed the highest CPA, mean heights, maximum heights, and minimum heights. *Dicksonia antarctica* contained the smallest CPA, mean heights, maximum heights, and minimum heights. Some *Dicksonia antarctica* tree ferns were incorrectly predicted based on the height range of 0 to 48.67 m since the tallest tree ferns in this site were 4.9 m tall. *Acacia melanoxylon* and *Nothofagus cunninghamii* were intermediate in height between the *Eucalyptus obliqua* and *Dicksonia antarctica*.

3.3.2.2 Object-based classification accuracy assessment

In the four class analyses (for classes *Eucalyptus obliqua, Acacia melanoxylon, Dicksonia antarctica* and Bare ground and non-vegetated objects), the fused MNF and CHM dataset yielded the highest accuracy (OA = 78.26%, k = 0.65) followed by the fused HSI, CHM and Indices dataset (OA = 75.76%, k = 0.62) (Table 3.4). These accuracies showed that the fused MNF and CHM dataset could produce better classification accuracy than that of the fused HSI and CHM dataset even if the indices were added to the fused HSI and CHM dataset. The fused HSI and indices dataset (OA = 72%, k = 0.43) resulted in higher accuracies than the HSI dataset alone (OA = 71.70%, k = 0.43). In this study, the *kappa* (k) of the fused HSI and CHM dataset (OA = 69.70%, k = 0.52) had a stronger agreement than the HSI dataset alone. The MNF dataset alone depicted the lowest classification accuracy (OA = 68.75%, k = 0.44).

Classification accuracy for the five class analyses was lower than that of the comparable four class analyses. In contrast to the four class analyses, the fused HSI, CHM, and indices dataset (OA = 66.67%, k = 0.52) produced the highest classification accuracy, followed by the fused MNF and CHM dataset (OA = 66.04%, k = 0.49). In this case, the HSI dataset alone showed the lowest classification accuracy (OA = 59.02%, k = 0.30). All the species classified in this research were evergreen and might share their spectral signatures that caused the difference in the classification accuracies. Marrs and Ni-Meister (2019), in this regard, suggested that the contaminating effect of non-dominant species' spectral signatures and shrub understory or bare ground are likely reasons for different classification accuracies.

The *Eucalyptus obliqua* class portrayed the highest sensitivity, indicating a correct classification of its presence, and *Nothofagus cunninghamii* was the least correctly classified among the five classes (Table 3.5). For four classes, the random forest classifier could identify 100% of *Eucalyptus obliqua* with the HSI dataset alone or fused

with CHM and indices datasets, whereas the MNF dataset alone or fused with CHM dataset produced the sensitivity of 92.86% and 96% respectively. The random forest classifier could identify 75% of *Dicksonia antarctica* with the fused HSI, indices and CHM datasets followed by the HSI dataset alone with the sensitivity of 71.43%. The fused HSI and indices dataset produced the least sensitivity (33.33%) with *Dicksonia antarctica* that demonstrated the deployed indices did not contribute to the HSI dataset for its identification. Out of four classes, *Acacia melanoxylon* had the lowest sensitivity with all the combinations of datasets except for the fused MNF and CHM dataset. As to specificity, *Eucalyptus obliqua* could be identified for 100% of absences with the fused HSI, indices and CHM dataset as well as the fused HSI and CHM dataset, whereas the specificity of *Dicksonia antarctica* and *Acacia melanoxylon* ranged from 80% to 95.65% and 84.62% to 97.62% respectively.

Moreover, for *Eucalyptus obliqua,* the fused HSI, indices and CHM dataset produced 92.11% balance accuracy followed by the fused MNF and CHM dataset. The least balance accuracy was obtained with the HSI dataset alone (57.90%). Thus, as the higher percentage of the balance accuracy, the higher the classification accuracy, this study achieved the highest balance accuracy of *Eucalyptus obliqua* in all the datasets besides the HSI dataset, followed by *Dicksonia antarctica* and then *Acacia melanoxylon.* Even when five classes were used, *Eucalyptus obliqua* demonstrated its dominancy for sensitivity and balance accuracy ranging from 77.27% to 96.77% and 73.39% to 86.11% respectively. Although *Dicksonia antarctica* is a short understory species, this species produced better sensitivity and balance accuracy than did *Acacia melanoxylon* and *Nothofagus cunninghamii. Nothofagus cunninghamii* demonstrated the least sensitivity and balance accuracy.

Regarding the contribution of CHM to MNF and HSI datasets for object-based tree species classification (Figure 3.7), classification accuracies increased in all cases except for the HSI dataset for four classes. When the forests were classified into five classes, CHM contributed to the HSI dataset with an increase in accuracy of 5.08%. The highest contribution of CHM was to the MNF dataset for four classes, with an accuracy increase of 9.51%. If the indices were added to the HSI dataset, CHM contributed less than the MNF dataset. These outputs indicate that CHM does not contribute significantly to the HSI datasets if the indices are added for object-based classification, but contribute to the MNF dataset. Finally, the classification maps were generated based on four and five classes, using all six datasets as shown in Figure 3.8 and Figure 3.9.

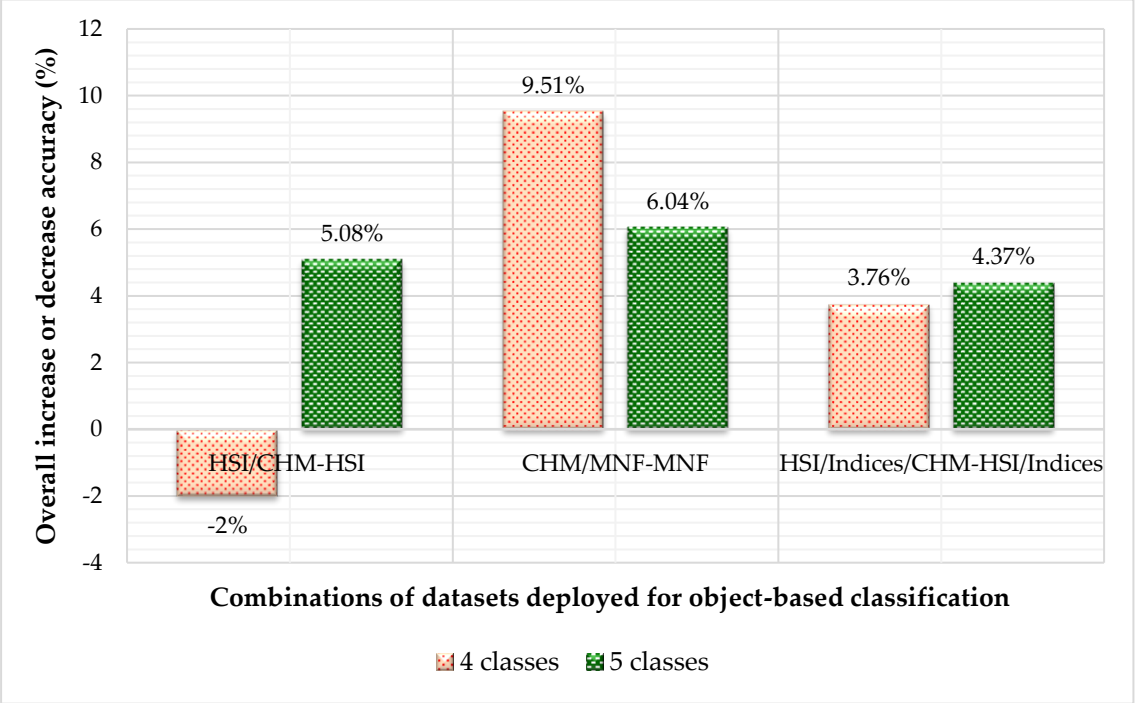


Figure 3.7 The contribution of the canopy height model (CHM) to the overall accuracy of the individual RS datasets. The contribution of CHM was assessed with or without CHM to minimum noise fraction (MNF), hyperspectral imagery (HSI), and fused hyperspectral and indices (HSI/Indices) datasets using object-based classification.

Table 3.4 Overall accuracy (OA) and kappa coefficient (*k*) based on object-based classification.

Datasets	4 classes		5 classes	
	OA (%)	***k***	**OA (%)**	***k***
MNF/CHM	78.26	0.6456	66.04	0.4909
MNF	68.75	0.4401	60.00	0.3351
HSI/CHM	69.70	0.5154	64.10	0.4479
HSI	71.70	0.4297	59.02	0.3049
HSI/Indices/CHM	75.76	0.6201	66.67	0.5173
HSI/Indices	72.00	0.4341	62.30	0.3895

Table 3.5 Species-wise sensitivity (Se), specificity (Sp) and Balance Accuracy (BA) based on object-based classification.

Class name	**MNF/CHM**			**MNF**			**HSI/CHM**			**HSI**			**HSI/Indices**			**HSI/ Indices/CHM**		
4 classes	**Se (%)**	**Sp (%)**	**BA (%)**	**Se (%)**	**Sp (%)**	**BA (%)**	**Se (%)**	**Sp (%)**	**BA (%)**	**Se (%)**	**Sp (%)**	**BA (%)**	**Se (%)**	**Sp (%)**	**BA (%)**	**Se (%)**	**Sp (%)**	**BA (%)**
A. melanoxylon	75	94.74	84.87	33.33	97.44	65.38	50.00	85.18	67.59	18.18	97.62	57.90	36.36	94.87	65.62	57.14	84.62	70.88
D. artarctica	62.50	92.11	77.30	50.0	85.00	67.50	66.67	80.00	73.33	71.43	95.65	83.54	33.33	95.45	64.39	75.00	86.21	80.60
E. obliqua	96.00	85.71	90.86	92.86	70.00	81.43	76.19	100	88.10	100	45.45	72.73	100	55.00	77.50	84.21	100	92.11
Bare ground and non-vegetated objects	20.00	95.12	57.56	0	95.56	47.78	66.67	100	83.33	0	100	50.00	0	97.87	48.94	66.67	100	83.33
5 classes																		
A. melanoxylon	57.14	93.48	75.31	28.57	95.35	61.96	33.33	78.79	56.06	16.67	95.92	56.29	40.00	92.16	66.08	50.00	86.67	68.33
D. artarctica	62.50	93.3	77.92	57.14	83.72	70.43	75.00	94.29	84.64	60.00	94.64	77.32	42.86	92.59	67.72	80.00	87.10	83.55
E. obliqua	96.00	75.00	85.50	92.31	58.33	75.32	77.27	82.35	79.81	96.77	50.00	73.39	96.77	66.67	81.72	83.33	88.89	86.11
N. cunninghamii	12.50	93.33	52.92	0	97.67	48.84	0	94.4	47.22	11.11	90.38	50.75	11.11	90.38	50.75	0	96.87	48.44
Bare ground and non-vegetated objects	20.00	95.83	57.92	0	100	50.0	75.0	100	87.50	0	100	50.0	0	100	50.00	66.67	96.97	81.82

(***Note***: Hyperspectral imagery (HSI), minimum noise fraction (MNF), canopy height model (CHM))

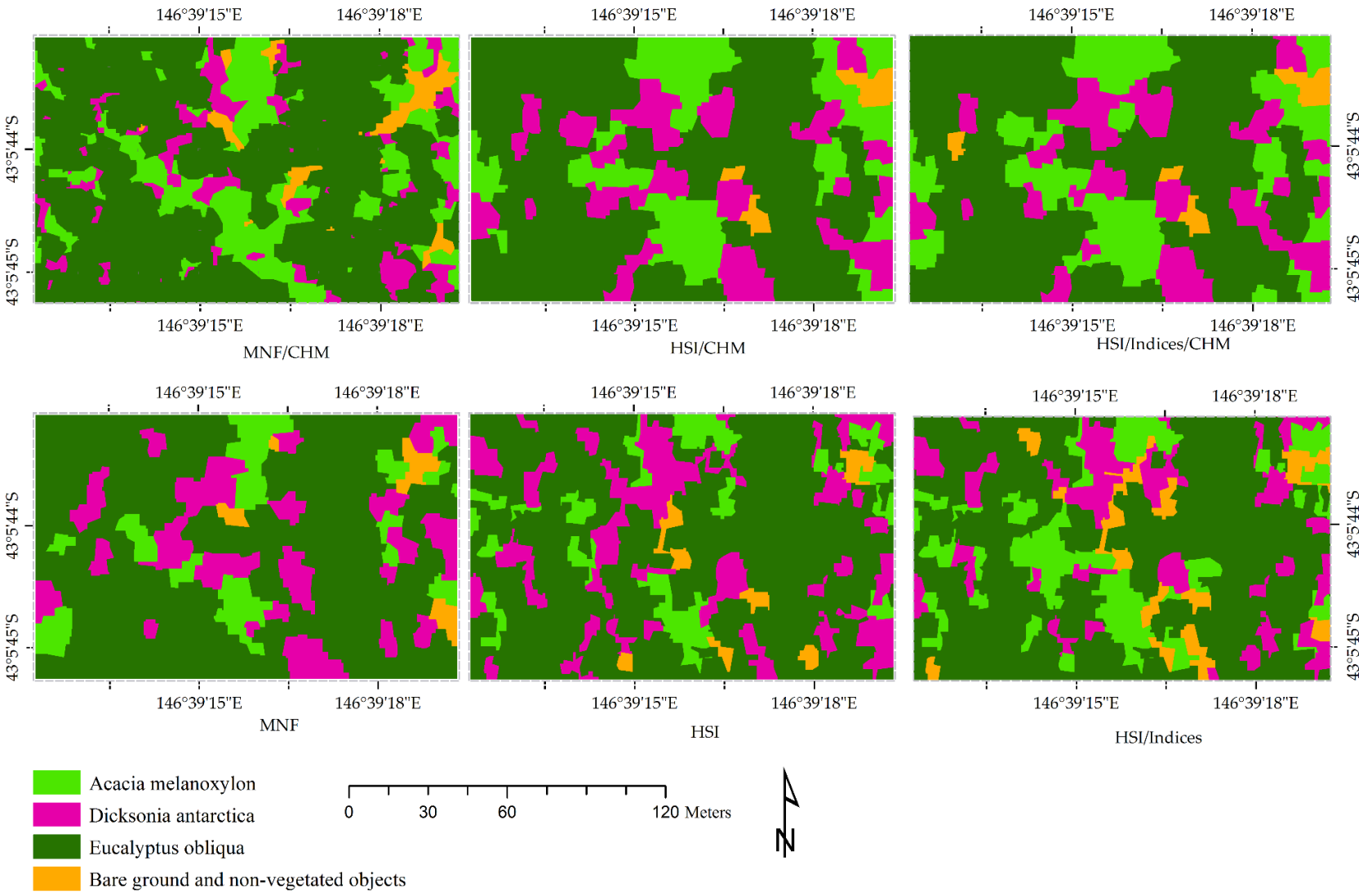


Figure 3.8 Object-based classification of four classes compared across six datasets.

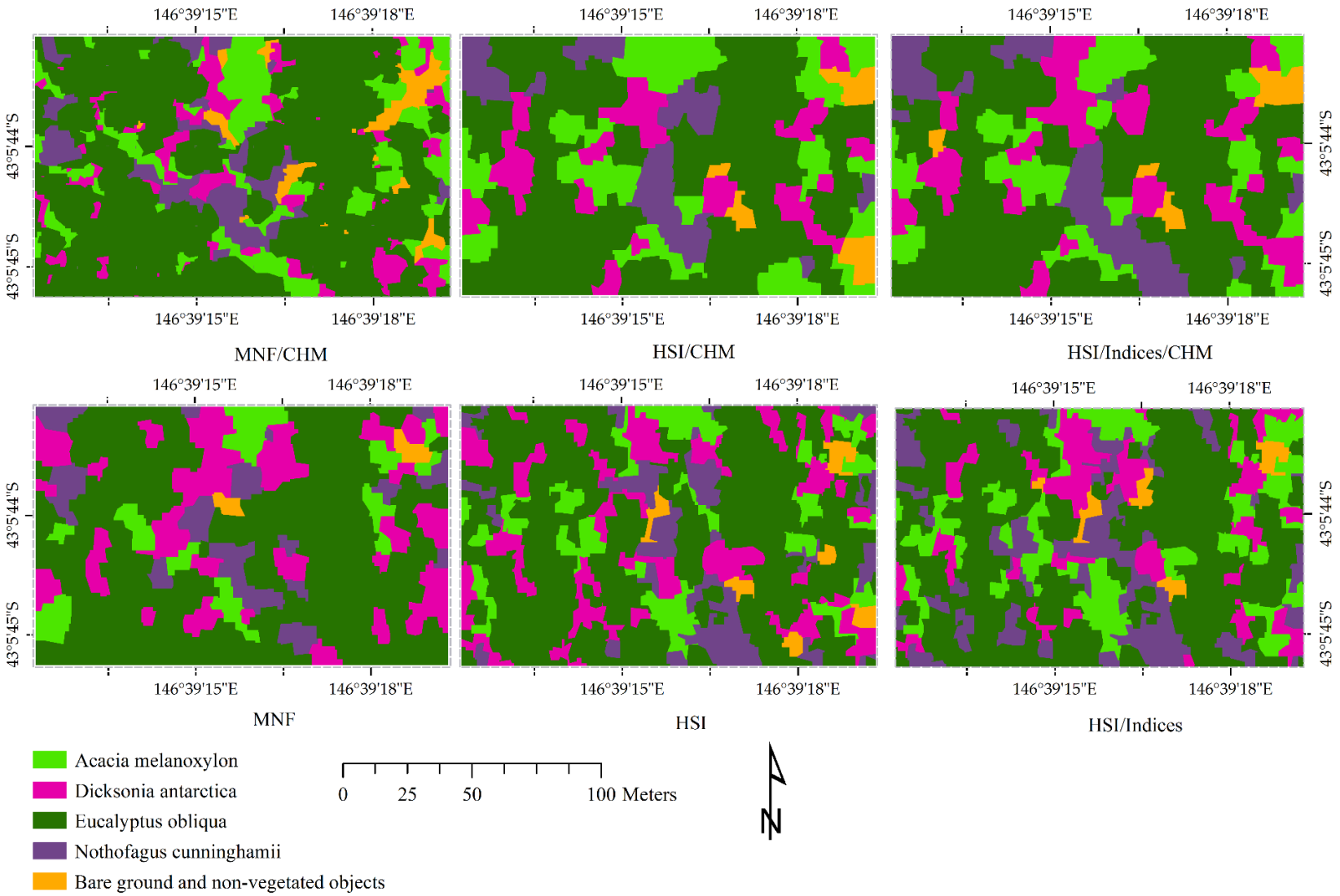


Figure 3.9 Object-based classification of five classes compared across six datasets.

3.4 Discussion

This work explored optimal segmentation of tree crowns using different combinations of remote sensing datasets, and classified dominant overstory and understory wet eucalypt forest species based on object-based classification.

For segmentation, HSI and MNF datasets were deployed with or without CHM to examine the performance and behavior of four datasets. Fusing MNF and CHM datasets, this study provided the highest segmentation accuracy (88.71%) with scale parameter, shape, and compactness of 10, 0.3 and 0.2 respectively. This was better than the results of Liu and Bo (2015) who obtained the accuracies of 61.60% with the fused MNF and CHM and 84.80% with the fused HSI and CHM for agriculture crops. Ke et al. (2010) evaluated segmentation quality using the relative area of an overlapped region to a reference (RA_{or}), the relative area of an overlapped region to a segmented object (RA_{os}), and position discrepancy (D_{sr}). They achieved RA_{or} values ranging from 0.6% to 32%, and RA_{os} values of 79% to 97% for the fused spectral and LiDAR-based segmentation. Most of the studies have not assessed accuracy to obtain optimal segmentation, and they have used 'trial-and-error' rule-based techniques (e.g. Ma et al. 2017; Moffett and Gorelick 2013) to visually select the final segments for classification purposes (Zhang et al. 2016). A few studies have segmented RS data fusing two or more RS datasets, for example, spectral/LiDAR (e.g., Ke et al. 2010), NDVI/CHM (Ruiz et al. 2016) and HSI/CHM and MNF/CHM (Liu and Bo 2015). In addition, studies have used only LiDAR data for tree crown segmentation (Dalponte et al. 2015; Devriendt et al. 2012; Lopatin et al. 2015; Zhang et al. 2016). Since it is unclear from the literature which combination of approaches is best, this study explored optimal segmentation procedures and parameters and applied the optimized segments for tree species classification. The best segmentation parameters in this study may not, however, be optimal for other applications (e.g. other forest types, agriculture or urban settings) because it depends on many factors like input layers, weights, color/shape ratio and scale factor (Ke et al. 2010). Nevertheless, the repeatability and transferability of the scale, shape, and compactness is still a challenging issue to determine optimal segmentation parameters and could be a topic of research to automate those parameters.

Previous studies have shown that object-based classification yields higher accuracy than pixel-based classification (e.g., Dian et al. 2016; Dorren et al. 2003; Zhang et al. 2016). Object-based classification provides a visually pleasing appearance of

classification (Duro et al. 2012; Ghosh et al. 2014). Comparing different datasets used in this study, the fused MNF and CHM dataset produced the highest overall accuracy among the six datasets when classifying vegetation into four classes. By contrast, when classifying vegetation into five classes, the fused HSI, CHM and indices dataset produced the highest classification accuracy among the six datasets tested. This result was comparable to those of many previous studies (e.g. Cho et al. 2012; Colgan et al. 2012; Féret and Asner 2012; Jones et al. 2010; Zhang 2014; Zhang and Qiu 2012). Thus, the fused HSI, CHM and indices dataset appears to be a good choice for tree species classification and mapping in temperate wet eucalypt forests. However, since greater accuracy was found for the classification of four (78.26%) rather than five (66.67%) vegetation classes, it appears that object-based classification may be unable to accurately classify a high number of species in diverse forest types. It should be noted here that the test site contained other plant species that this study did not attempt to identify.

Tree species classification is fundamental to plant species diversity monitoring and sustainable management. In this study, the dominant canopy species, *Eucalyptus obliqua,* had the highest sensitivity and balance accuracy, whereas the common understory tree *Nothofagus cunninghamii* had the lowest classification accuracy. *Dicksonia antarctica,* an understory tree fern, could be classified with a sensitivity of 75% using the fused HSI, CHM and indices dataset. These results highlight the advantages of the fused over single dataset for forest biodiversity and ecological studies. Although *Nothofagus cunninghamii* produced the least sensitivity and balance accuracy, it had a specificity of above 90%. Perhaps the spectral signature of this species was similar to that of the other species, something that could be tested with follow-up research.

The CHM contributed to the object-based tree species classification as shown in Figure 3.7. Adding the CHM to either the fused HSI and indices or MNF datasets increased the overall classification accuracy. Accuracy was improved by 9.51% for four classes fusing the MNF and CHM datasets compared to MNF dataset alone. Similarly, object-based classification accuracy was increased by 5.08% for five classes using the fused HSI and CHM dataset comparing with the HSI dataset alone. Here, CHM was able to contribute to the HSI dataset in dense and diversified forests indicating the usefulness of CHM for forest biodiversity monitoring and ecological studies. These outcomes were in line with Naidoo et al. (2012), Ke et al. (2010); Liu and Bo (2015). They

quantified the contribution of CHM for species classification and mentioned that understory tree species could be classified using the different height metrics.

Efforts were made to minimize the effects of potential error sources. However, all the error sources might not be eliminated completely. High dimensional HSI datasets did not accurately represent the specific species as other plant species had similar signatures. Although atmospheric corrections were made rigorously, the flight time of HSI data did not match exactly with MICROTOP II recording. The HSI, LiDAR data, and field data acquired were of different years. The time-gap between the acquisition of field data and HSI data was three years. A study carried out by Latifi et al. (2012a) had also suffered in their modeling accuracy from the three year time gap between the acquisition of reference data and the HSI datasets. Also, the size of training and test sets for the high dimensional data influences the performance of the random forest classifier, and consequently the classification accuracies.

3.5 Conclusions

Fused datasets rather than single remote sensing dataset were more robust for spatially discriminating and classifying wet eucalypt temperate forests at the species level using object-based tree species classification. Thus, the fused MNF and CHM datasets demonstrated the highest classification accuracy (OA = 78.26%, k = 0.65) when four classes were designated (*Eucalyptus obliqua, Acacia melanoxylon, Dicksonia artarctica,* and Bare ground and non-vegetated objects). When the number of classes was increased to five, the fused HSI, CHM and indices datasets produced slightly better results than the fused MNF and CHM datasets, although the overall accuracy was reduced (66.67%). So, this study recommends the fused HSI, CHM and appropriate indices dataset for tree species classification and mapping of wet eucalypt temperate forests, assuming that it is preferable to include more species classes at the slight expense of some classification accuracy.

Considering individual species, this study could classify *Eucalyptus obliqua* with balance accuracy of up to 90.86% when four classes were included, and up to 86.11% for five classes. The important understory species *Dicksonia antarctica* was identified with a balance accuracy of up to 83.54% for four classes, and up to 84.64% for five classes. This result demonstrates that not only dominant overstory species can be identified with this approach, but also a large crowned understory species.

The results demonstrate that the inclusion of the CHM dataset can improve the tree crown segmentation and species classification accuracy. The approach developed in

this chapter is likely to be useful for forest managers and ecologists for planning and sustainable management of wet eucalypt forest vegetation, and the species maps can be used for monitoring species of interest. Thus, this approach may provide a useful complement to the traditional photogrammetric mapping technique in Tasmanian forests. Future research should focus on the algorithms for the automated segmentation of objects that could be superior to the 'trial-and-error' approach. This study was confined to a 1.6 ha forest plot where tree locations were mapped. Hence, this approach needs to be tested at different study sites using multispectral imagery and low-density discrete return LiDAR data to ensure the transferability of the conceptual workflow to broader scales and other remote sensing datasets.

Chapter 4

Using topographic attributes to predict the understory structure of a wet eucalypt forest

Abstract

Forest understory structure is an important component of forest ecosystems that affects forest-dwelling species, nutrient cycling, fire behaviour, biodiversity, and regeneration capacity in a forest. Using traditional approaches to predict understory vegetation structure is difficult, but remote sensing may overcome these difficulties. Light detection and ranging (LiDAR) datasets have the capacity to predict understory structure as well as providing multiple resolutions of digital terrain models (DTMs) from the same dataset. High-density airborne LiDAR data were used to derive the densities of three understory layers as response variables. Topographic attributes for predictor variables were derived from DTMs at grid cell resolutions of 1 m, 5 m, 10 m, 20 m, and 30 m. The canopy density layers of LiDAR returns from ≥2 to ≤10 m, >10 to ≤30 m and >30 to ≤50 m were used as proxies for the lower, middle and upper layers, respectively. This study used random forest regressions based on twelve topographic attributes and a geology variable to model the densities of three understory layers across a 21.88 km^2 area of wet eucalypt forest located at the Warra Long Term Ecological Research Supersite, Tasmania, Australia. The strongest models were those for the 30 m spatial resolution. These 30 m resolution models successfully predicted the highest accuracy with the mean value of root mean square errors (MRMSE) of 7.61% (R^2 = 0.82) for the upper layer and the lowest accuracy for the middle layer (MRMSE = 9.76%, R^2 = 0.77) using training dataset. The validation dataset produced the highest accuracy for the lower layer (RMSE = 8.97%) with the 30 m resolution dataset followed by the upper layer (RMSE = 11.55%) and then middle layer (RMSE = 13.69%). Although the model for the middle layer structure was not strongly predictive, it provided valuable information on the relationships of structure to the environment. Variable importance depended on spatial resolutions and understory structural layers, and among the topographic variables, geology produced the highest importance followed by solar radiation. This study illustrates the relationships of the topographic attributes with understory vegetation structure in a wet eucalypt forest. To expand the scope of the current study, the predictive power of the models could be tested on a larger geographical area using lower density LiDAR point clouds and multispectral remote sensing data, given the high cost of high-density airborne LiDAR datasets. The present study should be applicable for predicting fuel loads and biomass and assessing biological diversity. Thus, this study could be useful for foresters and ecologists contributing to the planning of sustainable forest management and biodiversity conservation.

Keywords: Airborne LiDAR; Topographic attributes; Understory structure; Random forest; Variable importance

4.1 Introduction

Forest understory vegetation is an essential component of forest ecosystems (Latifi et al. 2017) that provides wildlife habitat and influences fire behaviour, nutrient cycling, biodiversity and regeneration potential (e.g. Campbell et al. 2018; Eskelson et al. 2011; Lindenmayer et al. 1999; Simonson et al. 2014b; Suchar and Crookston 2010; Tuanmu et al. 2010; Wing et al. 2012). In temperate forest ecosystems, most of the plant biodiversity is contained within the understory vegetation layers (Weisberg et al. 2003). Understory vegetation also contributes to reducing soil erosion and maintaining soil structure (Suchar and Crookston 2010). Biomass in the understory vegetation is one of the components of the forest carbon pool that is not considered in most of the forest inventories (Temesgen et al. 2015). Moreover, the vertical structure of forests is a vital attribute that affects habitat quality, including foraging opportunities and breeding resources, for many forest animals (Camprodon and Brotons 2006). As a result, the vertical structure could be used to develop a quantitative indicator of biodiversity for the assessment of sustainable forest management (Ferris and Humphrey 1999) and modelling understory vegetation characteristics in forested landscapes has become crucial and clear (Latifi et al. 2017). In this context, the aim of this study is to predict understory structural properties that are relevant for managing a forested wet eucalypt ecosystem and can be of high societal relevance in many biodiversity rich and fire prone areas (Morsdorf et al. 2010).

The wet eucalypt forest types comprise a *Eucalyptus obliqua* overstory with a mixed multi-layered understory of rainforest and sclerophyllous species (Hickey 1994b; Koch et al. 2008; Neyland 2001) that have established following past disturbances, mostly wildfire or forest harvesting (Forestry Tasmania 2009a). Understories include wet sclerophyll (early-successional, disturbance-adapted) and rainforest (late-successional) species (van Galen et al. 2018). Tall wet eucalypt forests are also found in North Queensland, Northern New South Wales, and Victoria (Wood et al. 2015b). The Tasmanian wet forests are generally tall, typically ranging from 40 m to 90 m in maximum height (Forestry Tasmania 2009b). The forests have experienced a long history of wildfire, with several major fires since the 1850s. The forests also range from being intensively managed for timber production to protected forests (Hickey et al. 1999). Since the 1960s, clearfell, burn and sow with eucalypt seed with rotations of 4-10 decades has been the dominant silvicultural treatment in the wet eucalypt Tasmanian forests (Hickey et al. 2001; Hickey and Wilkinson 1999). Variable retention has also recently been used as an alternative to clearfelling in some lowland wet eucalypt

forests (Baker and Read 2011). A variety of research has been conducted in wet eucalypt Tasmanian forests, especially at the Warra Long Term Ecological Research Supersite (Brown et al. 2001). For example, Neyland et al. (2012) investigated three dominant understory communities: one wet sclerophyll and two rainforests. Aggregated retention, a form of variable retention is nowadays preferred to clearfelling in wet eucalypt forests in old-growth sites to maintain mature forest biodiversity and structural diversity at the coupe level (Baker and Read 2011; Neyland et al. 2012; Scott et al. 2012). Areas of primary forest that have not been impacted by past harvesting provide important habitat for plants and animals in production forest landscapes. Knowledge of the structure of such areas can provide information about their potential habitat conditions for biodiversity. Remote sensing could, therefore, help understand current stand structure in the wet eucalypt forest landscape of the Warra supersite. Such research into forest stand structure is a useful complement to current and planned ecological field studies (Hickey et al. 1999). In this regard, Riedler et al. (2015) recommended that complex wet eucalypt forests should be monitored and conserved to foster high biodiversity.

Understanding the interaction of understory structural development and topographic attributes can support sustainable forest management (Amiri et al. 2017; Tuanmu et al. 2010). Mapping of forest attributes using aerial photography is of limited use for understory vegetation structure, which is largely obscured by the canopy eucalypts. Forest field inventories (typically carried out for different purposes) can give direct measures of understory structure but these are costly and time-consuming. To date, understory vegetation cover prediction is difficult using traditional explanatory variables, e.g. leaf area index, basal area (Wing et al. 2012), so obtaining detailed information on understory vegetation across large geographical area is still limited and its mapping is challenging (Tuanmu et al. 2010; Wing et al. 2012). In this regard, the topography is a major component affecting vegetation species composition, structure and functions from place to place. Topography affects microclimate, drainage, soil formation, wildfire impact, etc., and these factors, in turn, influence the distribution of plant species and structural composition. Topographic attributes, therefore, have the potential to inform sustainable management of forests (Wang et al. 2015). However, relationships between topographic attributes and vegetation cover have rarely been used for the purposes of forest management and restoration aimed at the conservation of forest biodiversity (Fu et al. 2004).

Multiple images with different spatial resolutions or scales are required to examine understory layers of forests with complex structure, as shown by Kamal et al. (2014) who used six resolutions (0.5, 1, 2, 4, 8, and 10 m) to study the specific details of mangrove features. Optimum pixel sizes are required to fill the scale gap between satellite measurements and input requirements (Hong et al. 2009), particularly in the areas where forest management frequently changes the forest structure (Nijland et al. 2014). Similarly, other studies have also considered different resolutions and an appropriate resolution or scale (e.g., Bocedi et al. 2012; Ferro and Warner 2002; Moore et al. 1991; Sharma et al. 2016).

New remote sensing technology can provide synoptic views of large areas of forests with an opportunity for forest resource assessment at different scales (e.g. Baldeck et al. 2015; Hudak et al. 2008; Jenkins and Coops 2011; Pesonen et al. 2008; Tomppo et al. 2008). Traditional passive remote sensing techniques have been used to survey forest characteristics over large geographical areas at lower costs (e.g. Amiri et al. 2017; Jenkins and Coops 2011). Advances in stereo image matching have automated the extraction of dense photogrammetric point clouds from the collected aerial photographs to map forest structure (Filippelli et al. 2019). Photogrammetric techniques, however, do not provide the same level of penetration through the canopies as LiDAR, and lack representation of mid and understory to generate reliable maps of the bare-earth terrain and vertical structural layers of forest canopies (Wallace et al. 2016). Airborne LiDAR – an active RS technique has shown great potential to capture the three-dimensional (3D) structure of forests (e.g. Campbell et al. 2018; Clawges et al. 2008; Coops et al. 2007; Holmgren and Persson 2004; Lefsky et al. 2002; Lindberg et al. 2015). LiDAR can penetrate its pulses through gaps in vegetation canopies and register multiple returns representing both canopies and bare-earth terrain (Moudrý et al. 2019) and is an established tool for mapping vertical structural layers of forest canopies and for numerous forest structure attributes (Filippelli et al. 2019). This technique is a widely used, effective and proven technology for forest mapping and management, and can capture forest stand structure from which understory vegetation characteristics can be extracted (e.g. Amiri et al. 2017; Campbell et al. 2018; Chang et al. 2016; Eskelson et al. 2011; Hamraz et al. 2017; Jakubowski et al. 2013; Luscombe et al. 2015; Martinuzzi et al. 2009; Murphy et al. 2011).

LiDAR sensors emit a laser beam at an operator-specified angle and receive the reflected energy. This technique determines the distance between the sensor to a target object using either discrete return or full-waveform systems. Except in forests with

unusually dense canopies (Goodwin et al. 2006; Hill and Broughton 2009; Morsdorf et al. 2010), airborne LiDAR systems can penetrate through the forest canopy to extract information on the density of understory layers (Bigdeli et al. 2018). In addition, highly accurate digital terrain models (DTMs) can be derived from LiDAR elevation datasets across a range of horizontal resolutions (Anderson et al. 2006; Jayathunga et al. 2018). Numerous topographic attributes can subsequently be derived from the DTMs for the purposes of forestry or hydrological studies at a variety of scales, e.g. slope, aspect, plan and profile curvatures (e.g. Franklin 1998; Wilson et al. 2007).

Many studies have attempted to use LiDAR data to characterise forest understory properties across a range of forest ecosystems. Campbell et al. (2018) quantified understory structure in subalpine fir forests in Utah with two dominant sagebrush forest types and highlighted LiDAR NRD (normalized relative point density) was superior to ORD (overall relative point density) in their modelling. They also assessed the effects of overstory vegetation density, canopy height and pulse density for characterizing forest understory structure. Using discrete-return LiDAR in both of the conifer and deciduous forest stands in Germany, Latifi et al. (2016) attempted to explore understory and overstory stand layers using LiDAR metrics and showed the top and the bottom layers could be predicted better than the middle layers. Singh et al. (2015) investigated LiDAR for detecting and mapping the understory invasive species *Ligustrum sinense* (Chinese privet) in North Carolina and found that topography attributes combined with vegetation and overstory produced the highest accuracies. Their study contributed to the implementation of management plans for conserving native biodiversity. In southern Spain, Simonson et al. (2014b) identified that shrub-layer presence varied significantly due to the effects of forest management under *Quercus suber* canopy but found less capacity to characterise understories under *Quercus canariensis* canopies because of a lack of understory first returns. They emphasized the importance of investigating topographic as well as canopy effects on understories for fire risk and biodiversity conservation. (Nijland et al. 2014) modeled understory species useful for grizzly bear species habitat by combining LiDAR-derived terrain with forest canopy information. They found that multiple data types could model complex relationships, but linear models are not always suitable for this purpose. Using discrete return LiDAR data acquired from eastern England for temperate deciduous woodland, Hill and Broughton (2009) mapped the understory from the leaf-off last return applying the height threshold of 1 m to separate the ground vegetation layer. Martinuzzi et al. (2009) investigated the presence and absence of understory shrub cover using airborne discrete return LiDAR data and found 83%

accuracy of presence. Maltamo et al. (2005) used 30 m grid cells to determine structural characteristics of heterogeneous boreal forests and examined the distributions of the canopy height density. They found that the number of understory trees could be computed as the proportion of hits above different height quantiles.

In Mountain Ash forests in south-east Australia, Jaskierniak et al. (2011) used mixture models that produced canopy profile indices of understory and overstory vegetation. Although LiDAR-based studies of south-east Australian forests found less predictability of understory structure due to a greater height and canopy closure (Lee and Lucas 2007), the tall open forests of Northern New South Wales provided good results. Goodwin et al. (2007) recommended the requirements of further research to clarify the applicability of LiDAR and high-resolution imagery metrics for dense Australian forests. Similarly, and Amiri et al. (2017); Eskelson et al. (2011); Suchar and Crookston (2010) reported that the robust cost-effective models to predict understory vegetation structure are important and necessary. Overall, all these studies show that it is possible to characterise understory structure in relatively open or simple forest types, but there are challenges in doing so in more complex forests, especially those with a dense canopy. The wet eucalypt forests of Tasmania represent a forest type that is moderately complex with moderate canopy density and is therefore amenable to testing the limits of the power of LiDAR to assess understory structure.

The objective of this study is to predict the understory density layers of a wet eucalypt forest using topographic attributes and geological data using high density airborne LiDAR point clouds to derive reference datasets for the densities of three canopy layers. This study was conducted in a complex forest landscape that is located at the Warra Supersite in southern Tasmania, Australia. This approach constitutes an improvement to the remote detection of understory vegetation. The findings from this study should be helpful for foresters and ecologists for the planning of biodiversity conservation, wildlife habitat and sustainable management of forests from the predicted information on understory vegetation structure.

4.2 Methods

4.2.1 Study area and datasets

This chapter utilised airborne LiDAR and ancillary data, i.e. geology map. A detailed description of the study area and datasets was given in Chapter 2. This study was conducted within a 5 km by 5 km forested area located at the Warra Supersite within the Wilderness World Heritage Area in Tasmania, Australia. The selected site is

topographically complex, comprising harvested sites, rivers, gullies and wet eucalypt forests. A 3.12 km^2 area was excluded from all sides of the study area to avoid edge effects and 2.79 km^2 was excluded comprising of previously harvested sites, roads, and rivers, and thus a 19.09 km^2 core native forest area was considered for this study.

This chapter followed an overall workflow as shown in Figure 4.1. Three response variables, i.e. canopy density layers, were derived directly from LiDAR point clouds and thirteen predictor variables, i.e. twelve topographic attributes extracted from DTMs and a geology vector data were deployed for random forest regression modelling. This study determined the optimal values of the two parameters of random forest modelling, *ntree* (number of trees to grow) and *mtry* (number of variables randomly sampled as candidates at each split), and examined the importance of variables. Five topographic attributes were then chosen to exemplify their effects on the generation of understory vegetation structure. The complete maps of the predicted understory layers were finally presented.

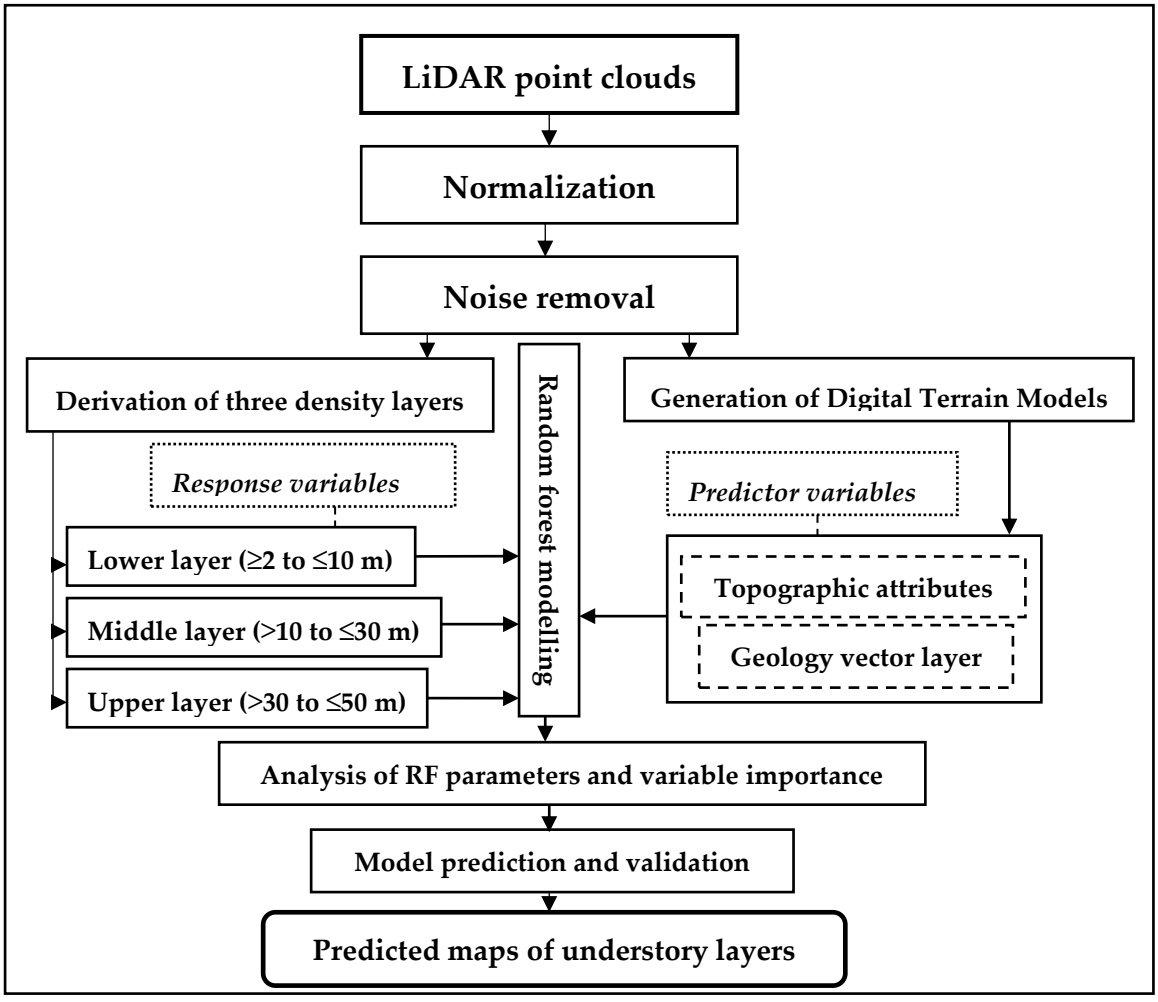


Figure 4.1 Research workflow diagram for the study.

4.2.2 Data pre-processing

LiDAR data were processed using LAStools (Academic version 171030). Firstly, LiDAR point clouds were classified into the ground and non-ground classes using *lasground*, then normalized using *lasheight*. The outlier points were removed from both above the canopy and below the ground level (Figure 4.2). To derive understory layers, relative height density rasters were generated using *lascanopy*. A 'relative height density raster' means counts of points falling into the height intervals are divided by the total number of points and scaled to a percentage within each grid cell. The DTMs were derived using *blast2dem*. Although the heights of understory layers depend on forest structure and composition, the landscape supports tall eucalypt species up to ~90 m tall. Based on field data from sample plots and field experience, most of the emergent overstory trees were above ~50 m in height, so the three understory layers were classified as lower layer (≥ 2 to ≤ 10 m), middle layer (> 10 to ≤ 30 m) and upper layer (> 30 to ≤ 50 m). On the other hand, the returns below 2 m (ground layer) were discarded from the analysis so that only representative canopy hits were considered, avoiding small shrubs and coarse woody debris (Ehlers et al. 2018; Packalén and Maltamo 2006; Wilkes et al. 2016). Musk (2017) reported that a large proportion of forest understory lies between 2 m and 10 m above ground.

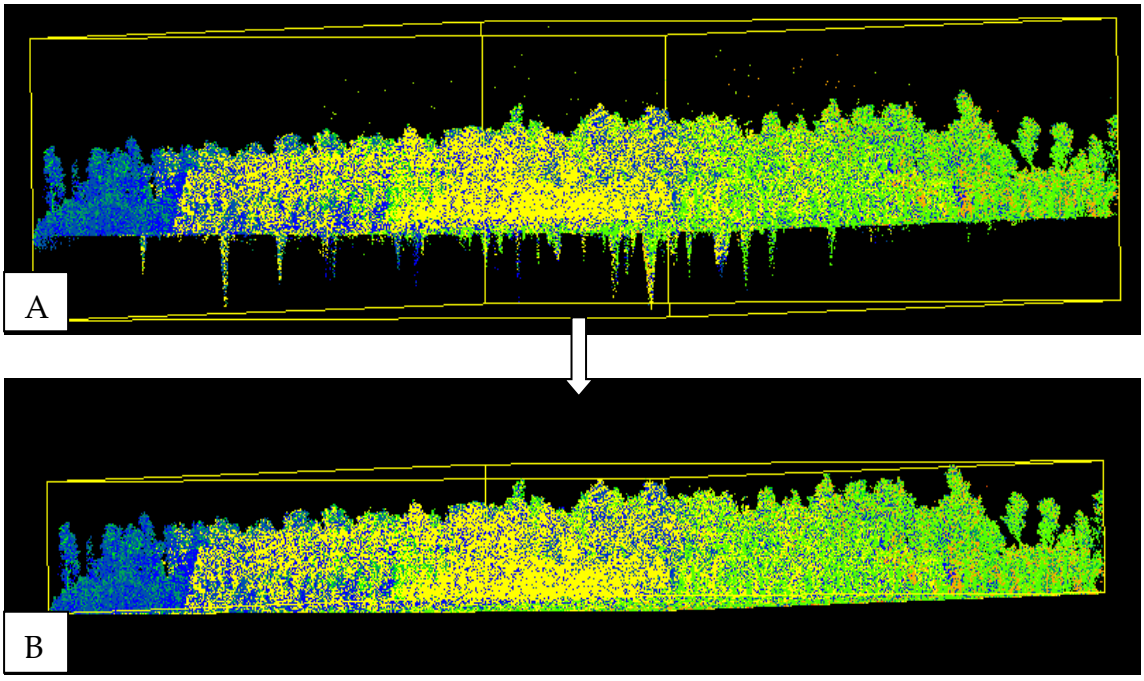


Figure 4.2 LiDAR point clouds of a tile based on Point Source ID in CloudCompare software: (A) showing outliers above and below ground level, and (B) showing de-noised LiDAR point clouds. Colours were displayed based on point ID ranging from 1 (blue) to 9 (red) with the default point size.

Previous studies have used LiDAR data for a variety of applications and their resolutions ranging from 0.25 m to 30 m (e.g. Gillin et al. 2015; Jaskierniak et al. 2011; Latifi et al. 2016; Latifi et al. 2012b; Thomas et al. 2017; Zimble et al. 2003). Jayathunga et al. (2018) considered the crown size of average trees to determine the grid size. The studies suggest that 10 m resolution of DTM would be sufficient for geomorphic and hydrologic modelling (please see Jenkins and Coops 2011; Murphy et al. 2011; Wang et al. 2011; Zhang and Montgomery 1994).

Examining different scales or resolutions to determine an appropriate resolution for this forest type was necessary, and as LiDAR data allowed us to generate different resolutions of DTMs, the resolutions of 1 m, 5 m, 10 m, 20 m, and 30 m were tested to predict the three understory layers using topographic attributes and a geology map. Maps of a digital terrain model of 30 m resolution derived from LiDAR data and geology types have been shown in Figure 4.3.

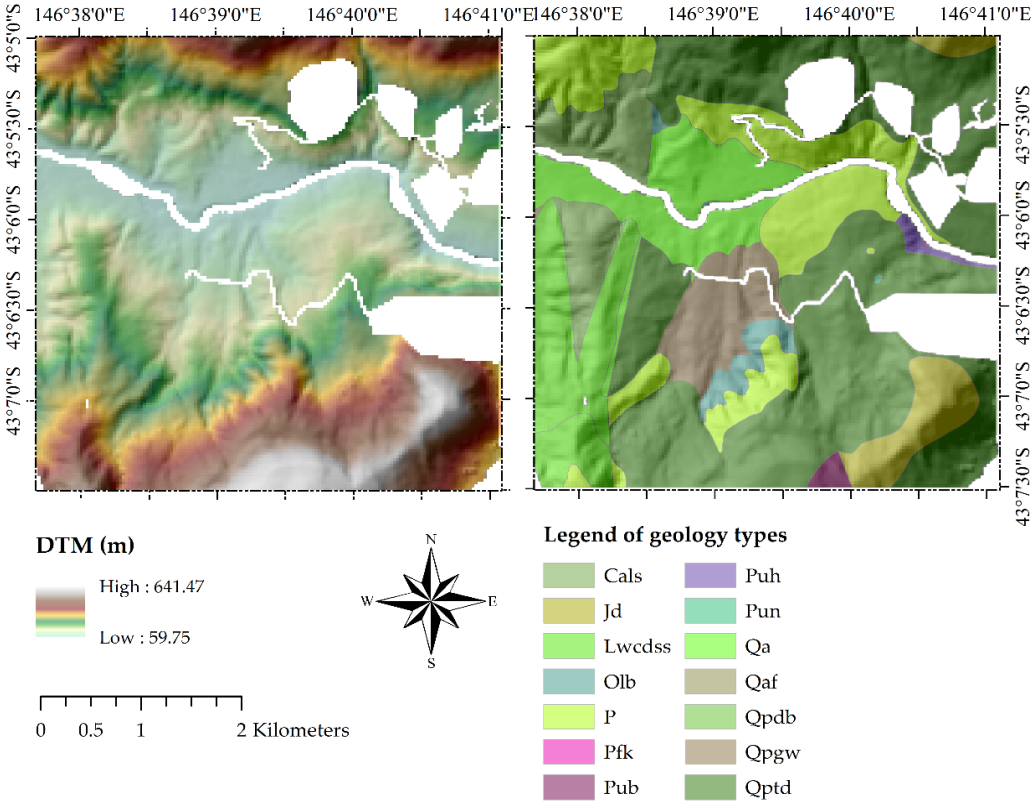


Figure 4.3 Maps of the study site sowing digital terrain model (left) of 30 m resolution derived from LiDAR data and geology types with a hillshade (right). In legends, Cals (Cambrian silicious sediments), Jd (Jurassic dolerite), Lwcdss (dolomite), Olb (Cambrian sediments), P (Permian sediments), Pfk (Carboniferous sediments), Pub (Carboniferous sediments), Puh (Carboniferous sediments), Pun (Carboniferous sediments), Qa (Quaternary alluvium), Qaf (Quaternary sediments), Qpdb (Quaternary sediments), Qpgw (Glacial tills), and Qptd (Dolorite talus).

4.2.3 Topographic attributes

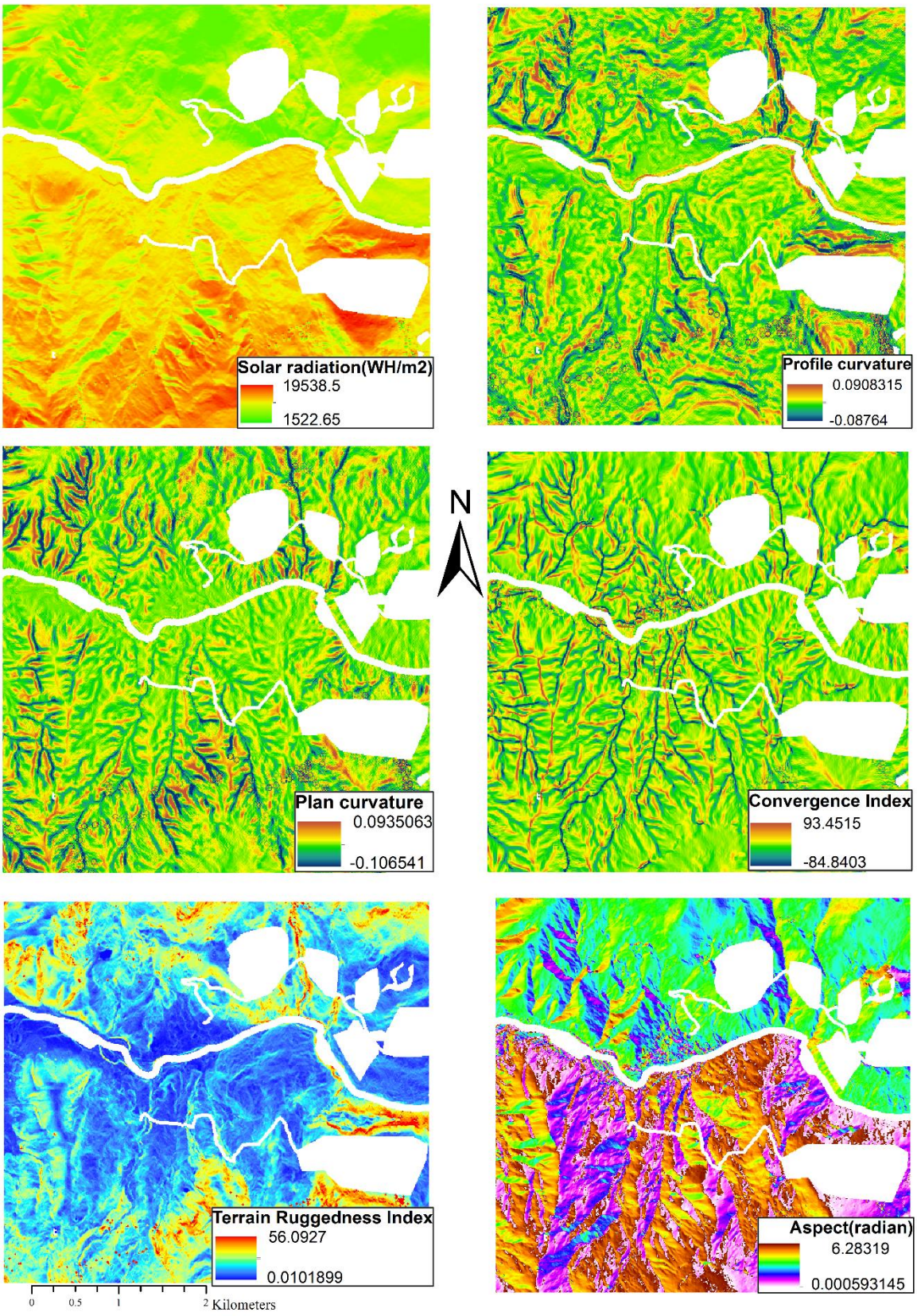
Topographic attributes include primary and secondary attributes (Table 4.1). Primary attributes are directly extracted from DTMs, such as slope, aspect, etc. Secondary attributes are calculated from the combinations of primary attributes and are mostly indices that characterize the spatial variability of specific processes occurring in the landscape, such as stream power index, terrain ruggedness index, etc. (Moore et al. 1991). Not all indices are likely to be closely related to understory densities (Wang et al. 2015), although some of the stream and catchment indices are thought to be relevant. Thus a subset of indices was selected that was expected to affect the forest species and their growth on site in different ways. For example, aspect and surface curvature are considered to impact understory vegetation (Wang et al. 2015).

To extract topographic attributes, the first step was to identify 'sinks and pits' using the method for LiDAR datasets by Wang and Liu (2006) cited in Thomas et al. (2016). Pits are often removed by filling them up to level the ground, however, that creates flat regions that are unsuitable for hydrological studies (Soille 2004) but are relevant to the vegetation growing on the terrain surface. Twelve topographic attributes extracted were slope, aspect, plan curvature, profile curvature, SAGA (System for Automated Geoscientific Analyses) wetness index, stream power index, terrain ruggedness index, terrain position index, solar radiation, convergence index, LS factor and catchment, from the same DTM in SAGA-GIS (Conrad et al. 2015). All the rasters of topographic attributes (Figure 4.4) were extracted at the same resolutions as the understory density rasters. A uniform resolution is required for all variables (Deng et al. 2007), especially for ecological modelling purposes.

Because geological attributes are important to vegetation, rock types were extracted from the geological map of Tasmania (scale of 1:25,000). The geology was simplified. Pub, Puh, P and Pun, all representing Permian sedimentary rocks, were integrated into a single category "Permian undifferentiated (P)" to avoid ambiguity in classification. Ordovician limestones were combined into one category (Olb) because the individual types were rare; dolerite boulders (Qpdb) had only a few points, so were merged with dolerite talus (Qptd). Quaternary sediments (Qaf) were uncommon and merged into a similar group Qa (Quaternary alluvium).

Table 4.1 Topographic attributes and their description.

Topographic attributes	Description	References
Slope	The slope is calculated by fitting a plane to the eight neighbouring cells or by finding the maximum slope of the cells.	(Travis et al. 1975)
Aspect	The aspect of the cell is the orientation of the cell relative to the north.	(Travis et al. 1975)
Catchment area	The upstream area of each cell. All grid cells with a cell's catchment area are upstream and drain into it.	(Kiss 2004)
Profile curvature	The rate of change of slope in a downslope direction. It can be used as a proxy for acceleration and deceleration of water over the terrain.	(Wilson et al. 2007)
Plan curvature	The curvature of a contour drawn through the central pixel. It describes the rate of change of aspect in plan across the surface and may be useful in defining ridges, valleys and slopes along the side of these features. It can be used as a proxy for convergence and divergence of water.	(Wilson et al. 2007)
LS (slope length and steepness) factor	Calculated using the upslope contributing the area of each cell and the grid cell slope because that depends on the volume of inflow water per unit cell rather than the flow length in the above cell.	(Desmet and Govers 1996; Kinnell 2005)
Potential solar radiation ratio	The ratio of the potential solar radiation on a sloping surface to that on a horizontal surface.	(Moore et al. 1991)
Topographic Position Index	It provides an indication of whether any particular pixel forms part of a positive (e.g., crest) or negative (e.g., trough) feature of the surrounding terrain.	(Wilson et al. 2007)
Terrain Ruggedness Index	This index is relevant for habitat mapping studies and calculates the sum change in elevation between a grid cell and its eight neighbouring grid cells.	(Riley et al. 1999)
Stream Power Index	A measure of the erosive power of flowing water	(Jacoby et al. 2011)
Compound Topographic Index or topographic wetness index	A measure of soil moisture potential that combines contextual and site information, and used to identify potential locations of ephemeral gullies.	(Casalí et al. 2016; Gessler et al. 1995)
Convergence index	Calculated by averaging the bias of the slope directions of the adjacent cell from the direction of the central cell and subtracting 90 degrees.	(Kiss 2004)


Solar radiation(WH/m2)
19538.5
1522.65
Profile curvature
0.0908315
-0.08764
N
Plan curvature
0.0935063
-0.106541
Convergence Index
93.4515
-84.8403
Terrain Ruggedness Index
56.0927
0.0101899
Aspect(radian)
6.28319
0.000593145
0
0.5
1
2
Kilometers

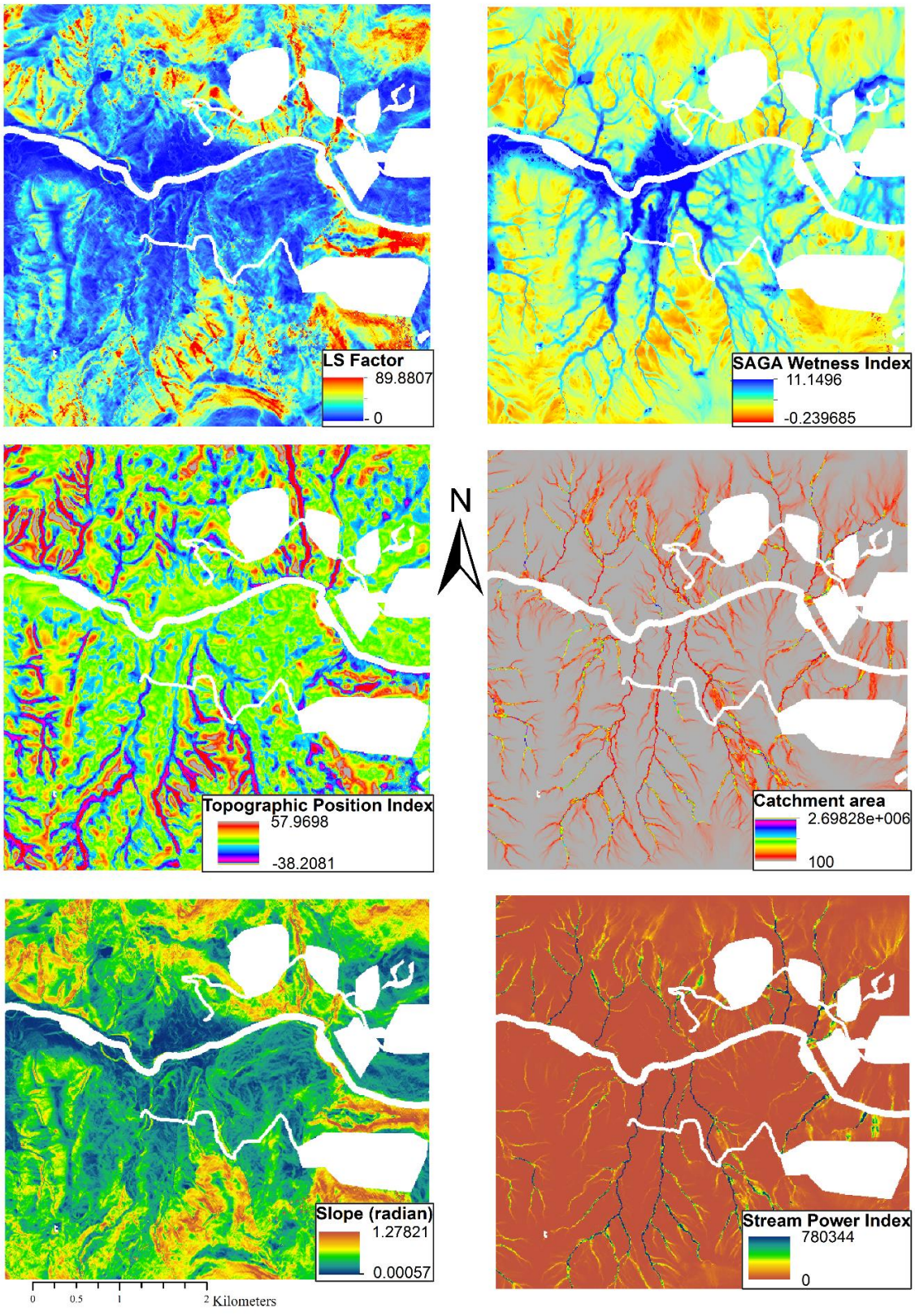


Figure 4.4 Maps of topographic attributes extracted from DTMs. All the predictor variables except for the geology map (See Figure 4.3) were extracted from DTM using SAGA-GIS software (Conrad et al. 2015). SAGA default parameters were used (Turner et al. 2018).

4.2.4 Random forest regression modelling

This study used random forest (RF) regression modelling to predict understory layers. A non-parametric machine learning algorithm developed by Breiman (2001), RF is extensively used in ecological modelling and remote sensing studies. The RF regression algorithm is a bagging technique, which employs recursive partitioning to divide the input data into many homogenous subsets called regression trees (*ntree*) and then averages the results of all trees. Each tree is independently grown to its maximum size based on bootstrap samples from the training dataset (approximately 67%) without pruning. In each tree, RF uses randomness in the regression process by selecting a random subset of variables (*mtry*) to determine the split at each node (Breiman 2001). In each tree, the ensemble predicts the data that are not in the tree (OOB (out of bag data), approximately 33%)), and by calculating the difference in the mean square errors between the OOB data and dataset used to grow the regression trees. The RF algorithm gives an error of prediction called the OOB error of estimate for each variable (Breiman 2001; Liaw and Wiener 2018; Prasad et al. 2006). This RF algorithm allows us to identify important predictor variables to predict understory vegetation structure (Martinuzzi et al. 2009). For the RF algorithm, variable importance is evaluated based on how much worse the prediction would be if the dataset for that variable were permuted randomly (Prasad et al. 2006), and this can be used in feature selection by determining the importance of each variable in the regression process (Freeman et al. 2015).

This study modelled the three understory layers as response variables based on the aforementioned twelve topographic attributes plus the geology vector data as predictor variables. The RF model training, prediction and validation processes were implemented in R (R Core Team 2017) using several libraries, including *randomForest* (Liaw and Wiener 2002) and *caret* (Kuhn 2017). To achieve stable results and test the sensitivity of the models, 100 iterations were run (Breiman 1996; Qian et al. 2016; van Galen et al. 2018; Yang et al. 2017b) for training and testing, and the predicted models were validated. For the selection of variables and final performance statistics, the average results were presented.

4.2.5 Sampling training and test datasets

This study used a random sampling method to investigate variation in the understory layers using topographic attributes and a geology dataset across the 5 km x 5 km study landscape. Altogether 10,000 random point locations were drawn (Criminisi et al. 2011;

Wood et al. 2011) from the study area and extracted values for each point for each of the variables. Approximately 1,700 points were deleted falling inside the previously masked areas, e.g., harvested sites, rivers, and roads. The remaining 8,300 sample points were divided into a model training dataset (50%) and a validation dataset (50%). Again, from the training dataset, 70% of sampled points were randomly drawn for the training and 30% for cross-validation (Abdel-Rahman et al. 2013; Kemppinen et al. 2018). The cross-validation was re-run 100 times to test the robustness of the model, and finally presented maps of understory layers.

4.2.6 Tuning parameters (*mtry* and *ntree*) for random forest modelling

Tuning the RF model, as implemented by the R package *randomForest*, only requires the user to make decisions about two tuning parameters – *mtry* and *ntree* (Cutler et al. 2012; Liaw and Wiener 2002; Tyralis and Papacharalampous 2017). The *mtry* parameter controls the number of predictor variables randomly sampled to determine each split. Higher values of *mtry* tend to work better in cases where only a few of the predictors contribute to the model and there are many predictors containing no useful information (Liaw and Wiener 2002; Prasad et al. 2006). In this study, the RF parameter optimisation was run for the 10 m resolution dataset only, because 10 m resolution lies in between the selected 5 resolutions, i.e. 1 m, 5 m, **10 m**, 20 m, and 30 m. Thirteen predictor variables were used, and thus this study considered five possible values for *mtry*: 3, 4, 6, 9 and 13. Another tuning parameter, *ntree*, controls the total number of independent trees. Predictions were made for the training dataset of the selected *ntree* from 200 to 10,000. The OOB error measure was plotted against a number of trees, with the *mtry* value of each model (Freeman et al. 2015). The *mtry* value of 13 (number of predictors) was applied (Abdel-Rahman et al. 2013; Di Lallo et al. 2017; Ghosh et al. 2014; Reese et al. 2014).

4.2.7 Variable importance

Variable importance is of interest for statistical analysis and prediction purposes to remove irrelevant variables (Genuer et al. 2015), and can be computed as the percent increase in mean square error (%IncMSE, also called the mean decrease in accuracy - MDA), or as the mean decrease in the Gini coefficient (MDG). %IncMSE was used, one of the most widely used scores of importance (Genuer et al. 2015) to evaluate the relevance of the variables for predicting understory layers using topographic variables (Louppe 2014; Sverdrup-Thygeson et al. 2016). %IncMSE is computed from permuting the test dataset. For each tree, the prediction error on the test dataset is recorded. Then

the same is done after permuting each predictor variable. The difference between the two is then averaged over all trees (Hastie et al. 2008; Liaw and Wiener 2018) and normalized by the standard deviation of the differences and if the standard deviation of the difference is zero for a variable, the division is not required (Liaw and Wiener 2018). Thus, the higher %IncMSE, the more important is the variable.

4.2.8 Partial dependence plots

Partial dependence plots (pdp) were used to produce a graphical representation of the marginal effect of predictor variables on the response variable (Friedman 2001). These plots are useful to explain the output of the predicted models like random forests (Greenwell 2017). The pdps consider the contributions of all the other predictor variables in addition to the particular predictor variable. In the pdps, the line is at the value of 0 indicates there is no impact of the predictor variables on the response variable. The higher the value, the greater the influence on the accurate prediction of the model. In this study, five predictor variables were selected with the higher variable importance scores and used the *pdp* package (Greenwell 2017) in R programming language to interpret predictor variables, i.e. topographic attributes and geology, and understory density layers as response variables.

4.2.9 Model accuracy assessment

This study used root mean square error (RMSE) (e.g. Gao et al. 2018; Kemppinen et al. 2018; Li et al. 2014; Rocha et al. 2018; Silva et al. 2017; Turner et al. 2018) to assess model training and testing applying 100 iterations, and averaged these to present model performance. Finally, the predicted models were validated with the independent dataset to test the robustness and stability of the model. RMSE produces an estimate of the standard deviation of the residuals (Alexander et al. 2015). The lower the RMSE, the better the model fit. The models were also tested using the coefficient of determination (R^2). The RF regression models were calculated using RMSE as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((\hat{y}_i - y_i)^2}$$

Where n is the number of validation plots, y_i is the reference value for the plot, i. $\hat{y}_i$ is the predicted value for plot i, and $\bar{y}$ is the mean of the reference variable.

4.3 Results

4.3.1 Tuning and selection of random forest parameters

As shown in Figure 4.5 to Figure 4.7, *ntree* from 200 to 10000, and *mtry* with 3, 4, 6, 9, and 13 were tested using the spatial resolution of 10 m. Three different understory density layers, i.e. lower layer, middle layer, and upper layer, were used to confirm the impact of *mtry* and *ntree* choice. The OOB errors of the selected *mtry* for the lower layer ranged from 0.01% for the *mtry* of 13 to 21.1% for 3 (Figure 4.5). For the middle layer, the OOB errors ranged from 0.02% for the *mtry* of 13 to 13.82% for *mtry* of 3 (Figure 4.6), and for the upper layer, OOB errors ranged from 0.05% for *mtry* of 13 to 14.78% for 3 (Figure 4.7). Thus, the *mtry* value of 13 consistently yielded the lowest OOB errors, whereas the default *mtry* (4) is one-third of total predictor variables and that produced the higher OOB error. The difference between the default and optimised *mtry* values justified a detailed description of this optimisation process.

Considering different values of *ntree* with a *mtry* set at 13, the upper layer provided the OOB error of 0.06% at the *ntree* of ≥ 400 and 0.05% at the *ntree* from ≥ 2200 to 10,000. Similarly, the OOB error for the middle layer produced 0.03% for *ntree* values of < 500, and then 0.02% from ≥ 500 to 10,000. Thus, this study showed that the *ntree* parameter did not differ significantly with increasing *ntree* of >500, as usually considered default in the RF regression modelling (e.g. Turner et al. 2018). Thus, the combination of *mtry* of 13 and *ntree* of 500 was chosen to use for subsequent RF modelling purposes.

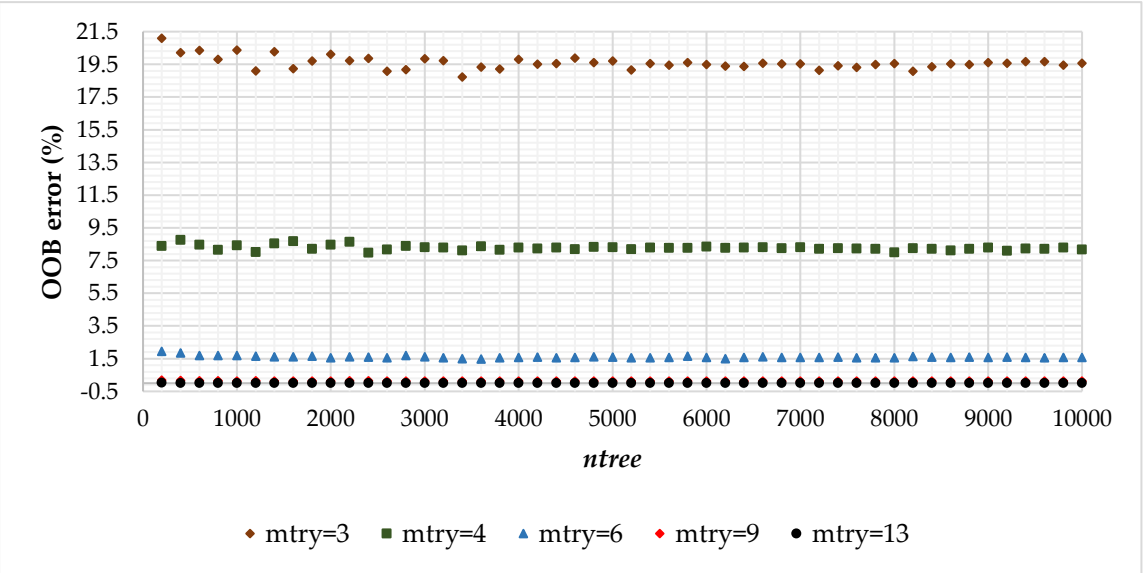


Figure 4.5 OOB error rates for different values of *mtry* and *ntree* for the lower layer of canopy density.

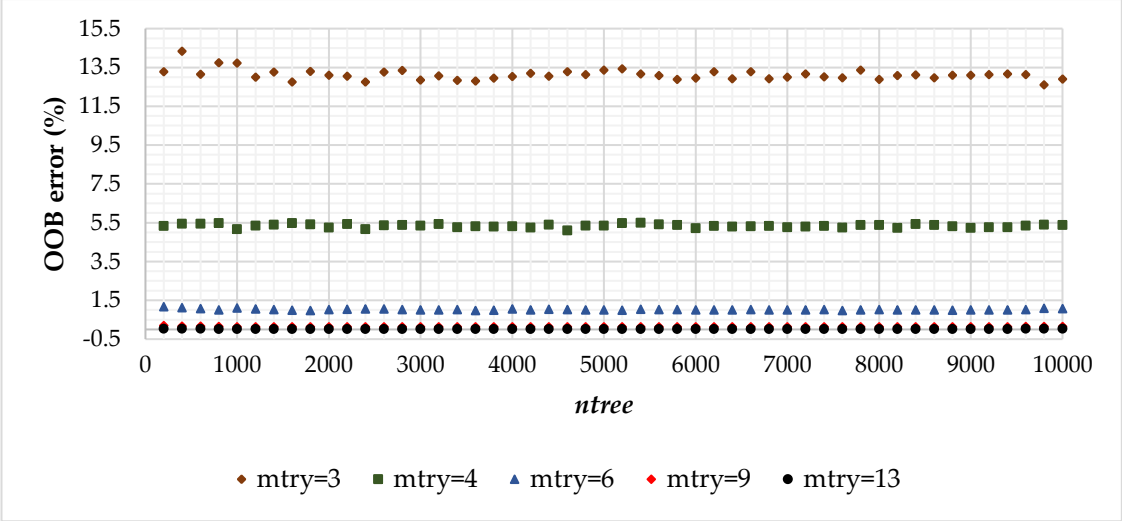


Figure 4.6 OOB error rates for different values of *mtry* and *ntree* for the middle layer of canopy density.

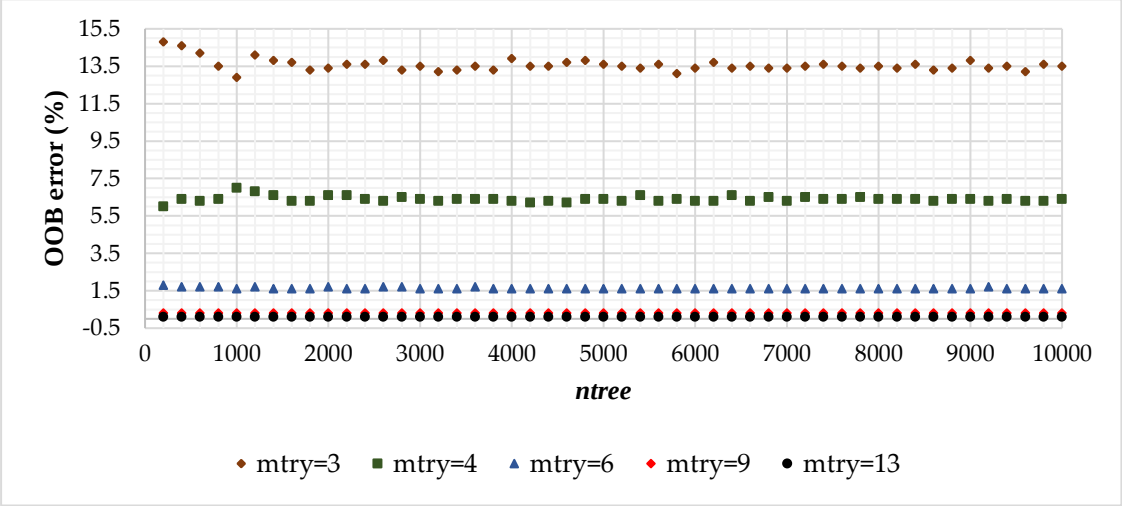


Figure 4.7 OOB error rates for different values of *mtry* and *ntree* for the upper layer of canopy density.

4.3.2 Random forest model prediction and cross-validation

Using the selected RF parameters, i.e. *mtry* (13) and *ntree* (500), for predicting three understory density layers, the model performance of 1 m, 5 m, 10 m, 20 m and 30 m resolutions was examined, 70% of training dataset was used and cross-validated with 30% of the test dataset using 100 iterations, and the averaged accuracies were presented (Figure 4.8). The predictive power for the upper layer ranged the MRMSE values from 7.61% (R^2 = 0.82) to 17.82% (R^2 = 0.64). The 30 m resolution dataset produced the lowest MRMSE value of 7.61% for the upper layer followed by the lower layer (MRMSE = 8.76%). The 1 m resolution dataset had the lowest predictive power in comparison with the other resolutions for all three structural layers. The predicted

results showed that the model accuracies increased with a coarser spatial resolution for all three layers, but this improvement was most marked for the lower layer.

The middle layer had the lowest predictive power among the three layers in the cases of 20 m and 30 m resolutions, but that produced better accuracies than the lower layer in the cases of 1 m, 5 m, and 10 m. There is only 1% of the difference between middle and lower, so they are comparable for 30 m pixel sizes. Thus, the 30 m resolution model was the best performing for all three canopy layers, with reasonable predictive power (MRMSE < 10%) for all three vegetation layers.

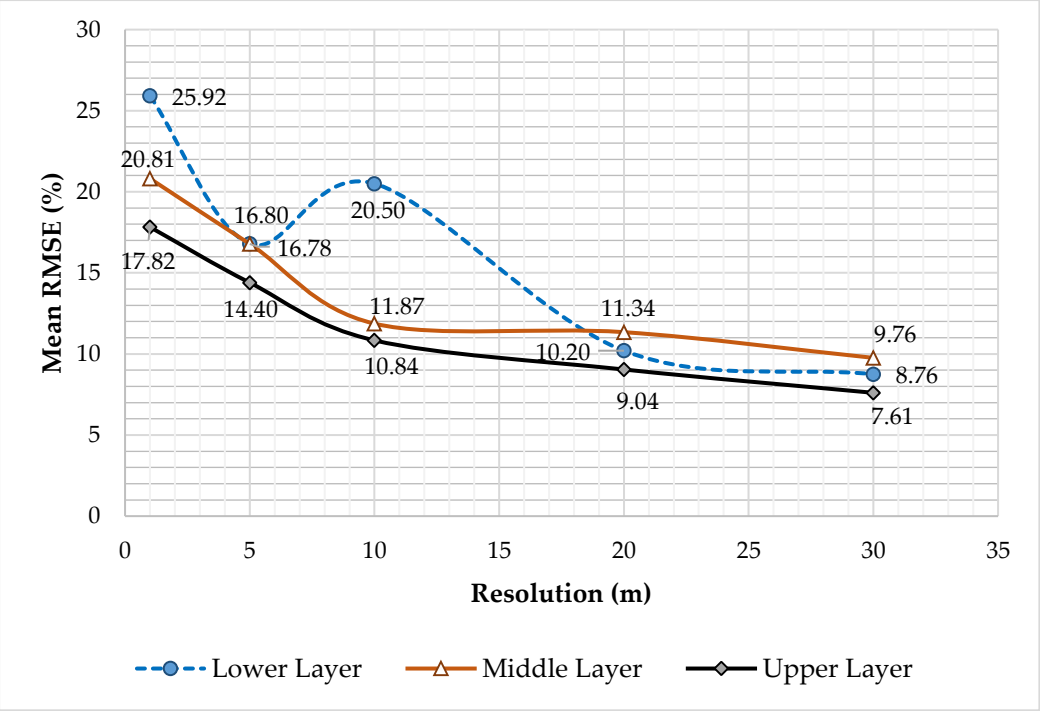


Figure 4.8 Mean values of root mean square errors of models using the selected resolutions for three density layers. Lower RMSE indicates a better model fit. As the unit of response variables (understory density layers) are in percentage, RMSE is measured in percentage.

4.3.3 Variable importance

Variable importance scores, i.e. percent increase in mean square error (%IncMSE), of different topographic attributes demonstrated different performance in predicting understory density layers (i.e. lower, middle and upper layers) (Figure 4.9 - Figure 4.13). Comparing the variable importance scores with %IncMSE values for the upper layer, geology provided the highest variable importance score for four resolutions (i.e. 5 m, 10 m, 20 m and 30 m), whereas terrain position index (TPI) produced the highest %IncMSE (37.36%) for 1 m resolution. For the middle layer, solar radiation produced

the highest %IncMSE of 19.58%, 57.37% and 89.66% for 5 m, 20 m and 30 m resolutions respectively, whereas plan curvature provided the highest %IncMSE value of 13.09% for 1 m, and geology (%IncMSE = 56.01%) for the 10 m resolution. With the variables for the lower layer, geology provided the highest variable importance score with %IncMSE values of 36.05%, 45.99% and 67.50% respectively for 5 m, 20 m and 30 m resolutions, whereas plan curvature provided the highest variable importance score with %IncMSE value of 21.56% for 1 m, and solar radiation (%IncMSE = 39.58%) for 10 m. Thus, geology provided the best overall result for the upper and lower layers, and solar radiation had the best performance for the middle layer in a wet eucalypt forest. The relative importance of other variables varied somewhat depending on the structural layer and resolution. So, these scenarios showed that spatial resolutions had a significant impact on the variable importance scores of the input variables.

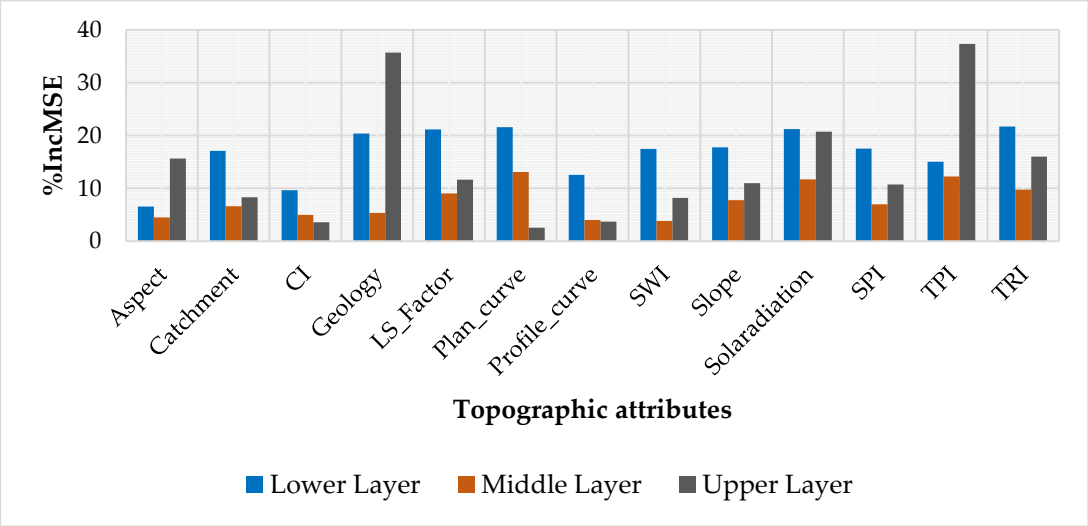


Figure 4.9 Variable importance scores for 1 m resolution.

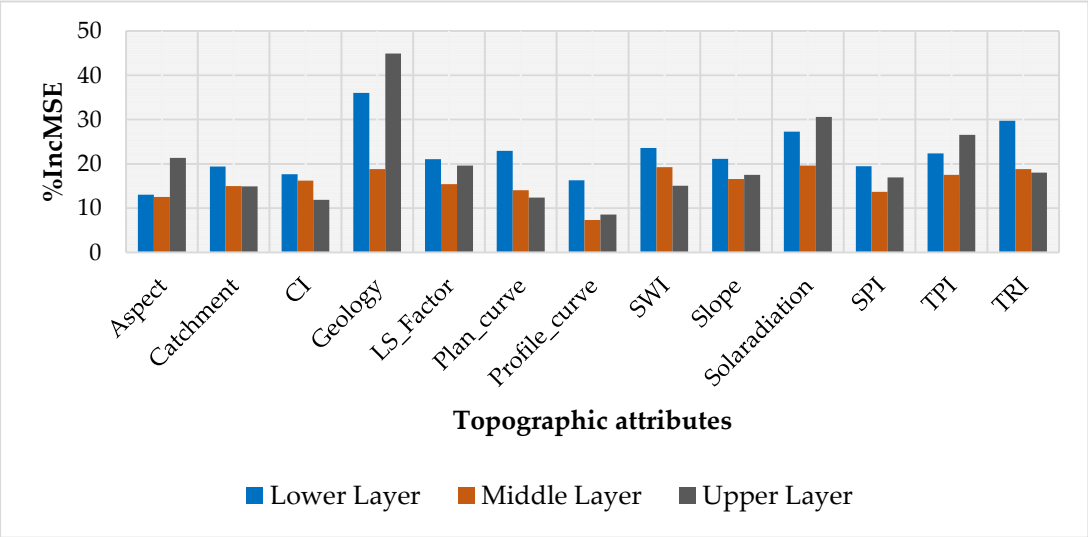


Figure 4.10 Variable importance scores for 5 m resolution.

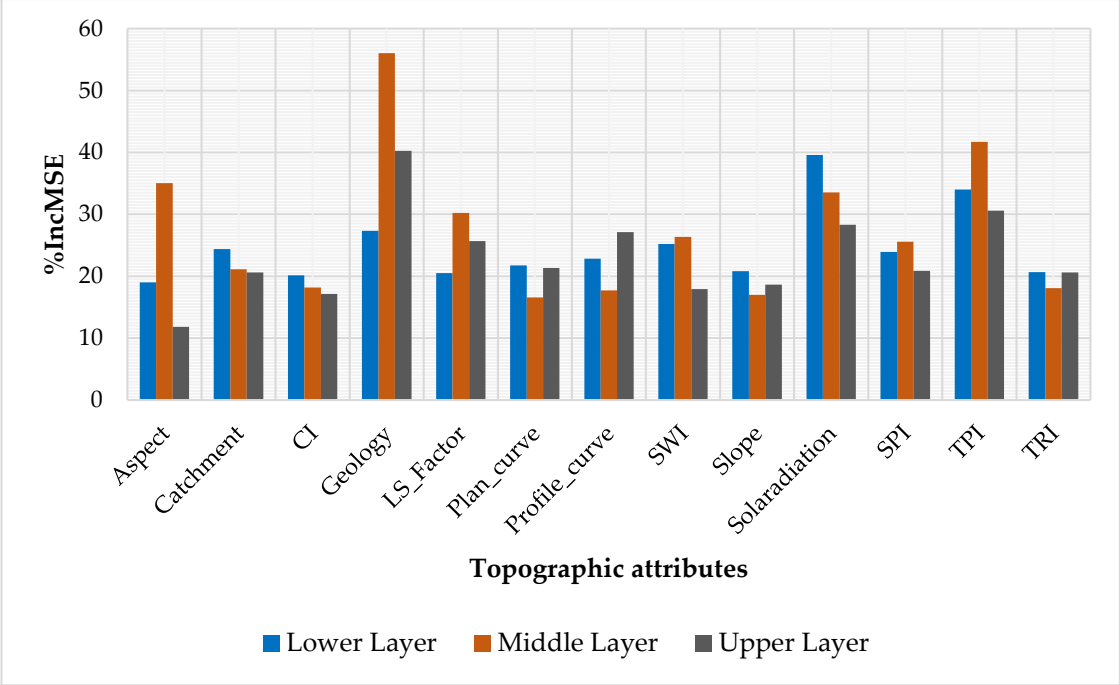


Figure 4.11 Variable importance scores for 10 m resolution.

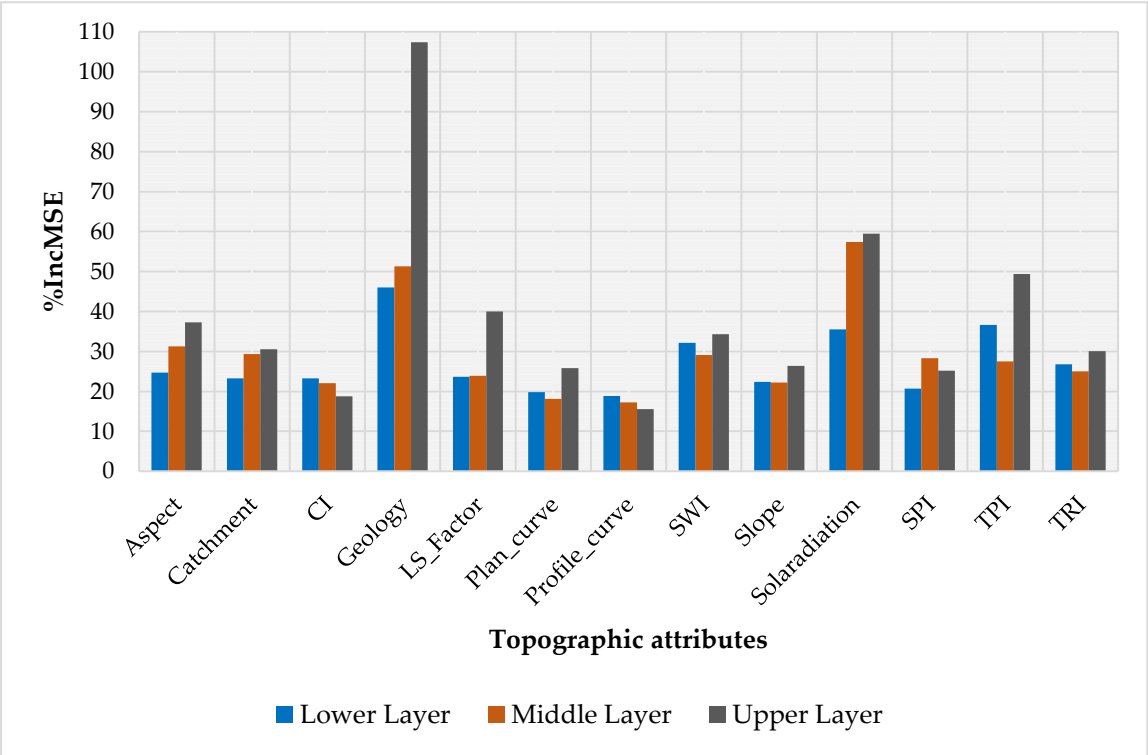


Figure 4.12 Variable importance scores for 20 m resolution.

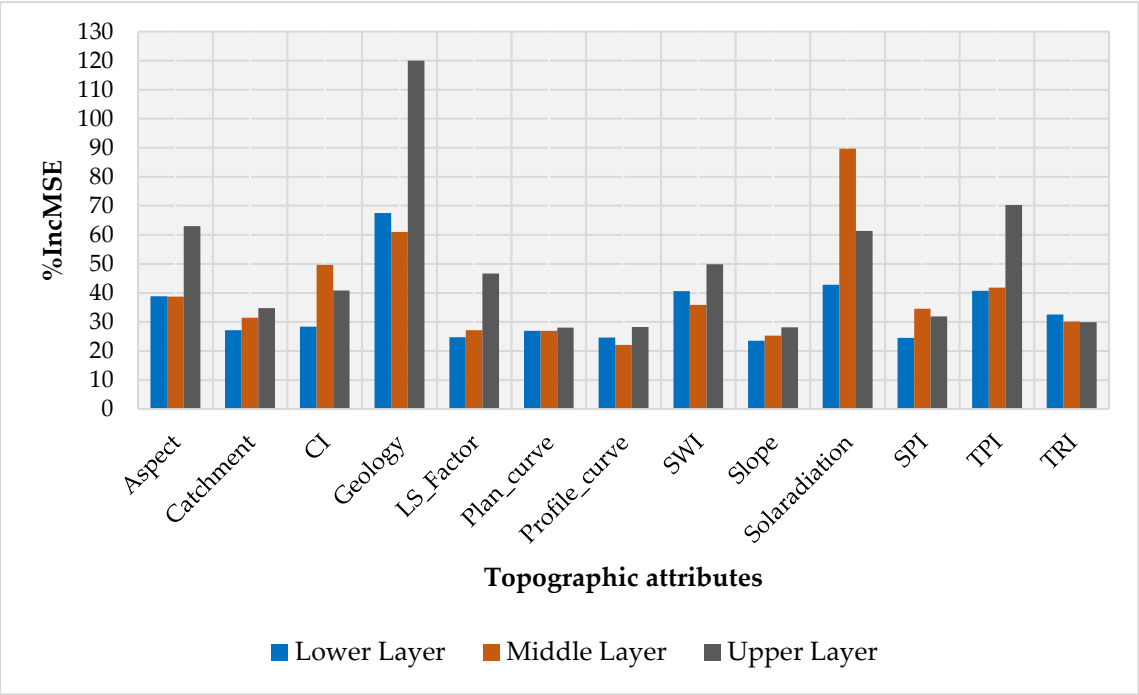


Figure 4.13 Variable importance scores for 30 m resolution.

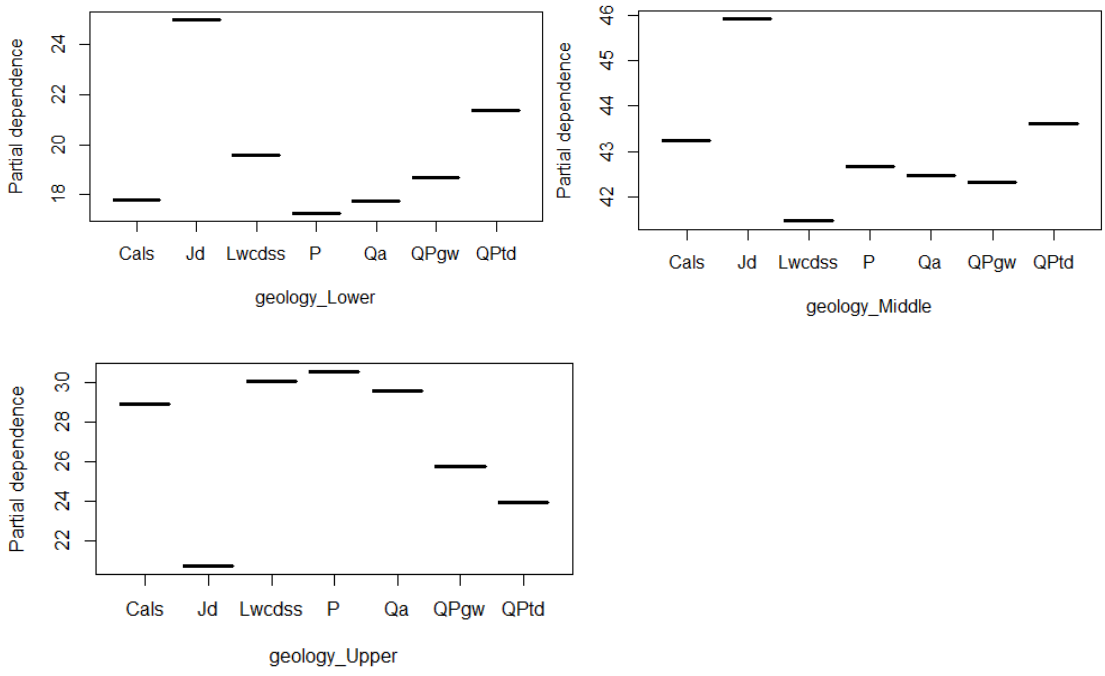
4.3.4 Assessment of partial dependence plots

Because the 30 m resolution model produced the best results, this dataset was selected for the purpose of displaying partial dependence plots. Based on results of the RF modelling with all the 13 predictor variables to model three understory density layers, this study focused on the best five predictor variables (i.e. geology, solar radiation, TPI, aspect, and SWI) for presenting partial dependence plots and considered the higher importance scores to measure the marginal effect of these five predictor variables on the understory layer models (Figure 4.14).

All the five variables had different influences and thus had different importance on the three understory canopy layers (Figure 4.14). For example, Jurassic dolerite (Jd) of the geology layer demonstrated the highest influence for the lower layer and middle layer of the models with pdp values of 26% and 46% respectively, yet had the lowest influence on the upper layer model. The Permian sediments (P) of geology showed the highest influence (pdp value of > 30%) for the upper layer model, whereas that provided the lowest influence (pdp value of < 10%) for the lower layer model. Another variable, the influence trend of solar radiation dropped from 5,000 to ≤ 15,000 (pdp values from 21.5% to approximately 5%) and after that reached the highest influence with the pdp value of > 24% at the solar radiation value of ≥ 25000 for the lower layer model. With the middle layer model, the pdp value showed the highest (> 46%) at the solar radiation value of 5,000 and then decreased to the pdp value of 40% at around

18,000. For the upper layer model, the solar radiation had the lowest influence (pdp value < 10%) at 5,000 and reached to the highest influence with pdp value of 28% at 18,000, and then dropped to around 25% at ≥ 25000 of solar radiation. These results indicated that solar radiation had a higher influence on the upper layer of the canopy at the mid-values of around 17000 than the lower and the higher solar radiation. For the middle layer, a lower value of solar radiation had high influence, whereas a higher value of solar radiation did not have a significant influence. Similarly, for the lower layer of the canopy, the higher the solar radiation values, the more its influence on the understory.

Similarly, the topographic position index (TPI) for the lower and middle layer models showed a similar trend of influence from -20 to 0. The TPI value at -20 had the highest influence (pdp value of 48%) for the middle layer model followed by the lower layer model (pdp value of 26%). Conversely, TPI value at -20 had the lowest influence (pdp value of < 10%) on the upper layer model and the highest influence (pdp value of >28%) at the TPI value of 9. The West-North (0 to 2^c, unit in radians) and East-North (4 to 6^c) aspects had the highest influence for the lower and upper layer models with the pdp values of ≥ 21.5% and 30% respectively, whereas the East-South aspect (2^c) had the highest influence on the middle layer model. The influence of SAGA wetness index (SWI) increased with the increase in the SWI value from 2 to 6 (pdp values ranged from > 20% to 28%), and then that stabilized with the TWI value at ≥ 6 for the upper layer model. This contrasts with the lower and middle layer models for which low TWI values had maximum influence.



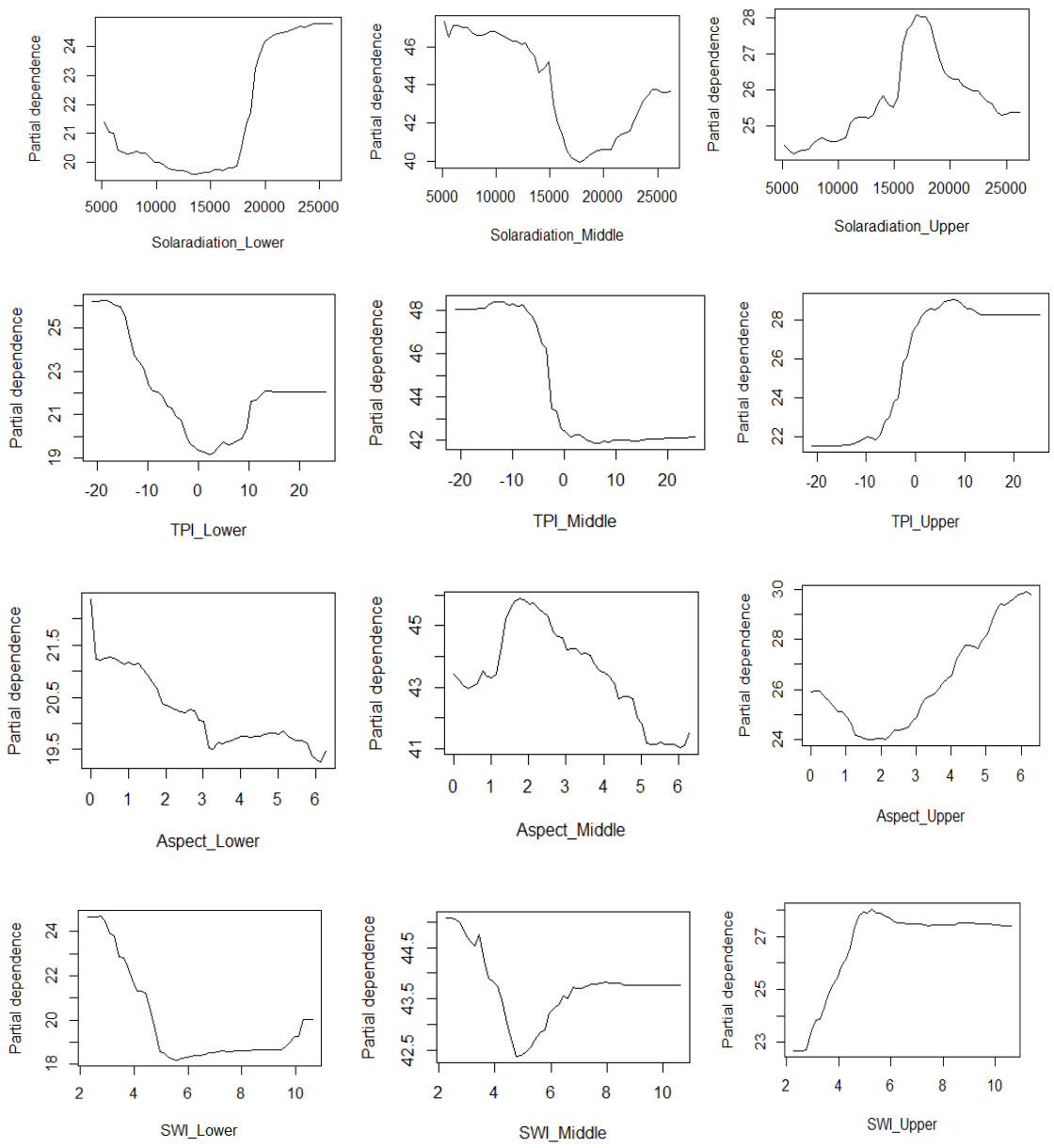


Figure 4.14 Partial dependence plots of geology, solar radiation, terrain position index (TPI), aspect and SAGA wetness index (SWI) of the lower layer, middle layer, and an upper layer of forest structure. Higher partial dependence values indicate a greater influence of the particular variable on the model. The geocodes used are Cambrian silicious sediments (Cals), Jurassic dolerite (Jd), dolomite (Lwcdss), Permian sediments (P), Quaternary alluvium (Qa), glacial tills (Qpgw) and dolorite talus (Qptd). The unit of aspect is in radians (c).

4.3.5 Random forest model validation

All the available training samples were split into 70% for training and 30% for cross-validation, and the RF models were trained and cross-validated using 100 iterations.

Using 50% of the independent validation dataset, model accuracies were validated (Figure 4.15) and predictive maps presented (Figure 4.16). The lower layer of the canopy densities with the 30 m resolution yielded the lowest validation RMSE of 8.97% (R^2 = 0.37), providing the best model fit, followed by the 20 m resolution with the RMSE of 11% (R^2 = 0.27). The middle layer for the 30 m dataset had higher RMSE values of 13.69%, so it provided the weakest predictive power for all five resolutions except for the 10 m resolution. The validated models for the lower layer produced the highest model accuracies for four resolutions, with the exception being the 10 m resolution where the upper layer provided the highest accuracy, followed by the middle layer. The spatial resolution of 1 m had the highest RMSE (32.97%) for the middle layer followed by the upper layer (RMSE = 27.27%), and the lower layer of the understory provided the lowest RMSE (25.27%) indicating that the RF model could predict better for the lower layer than the middle and upper layers. However, the 1 m resolution of the dataset could not predict well in comparison to the other four resolutions. The validation results showed that the model accuracies increased with an increase in pixel size indicating the impacts of resolutions on the prediction of understory density layers in a wet eucalypt forest.

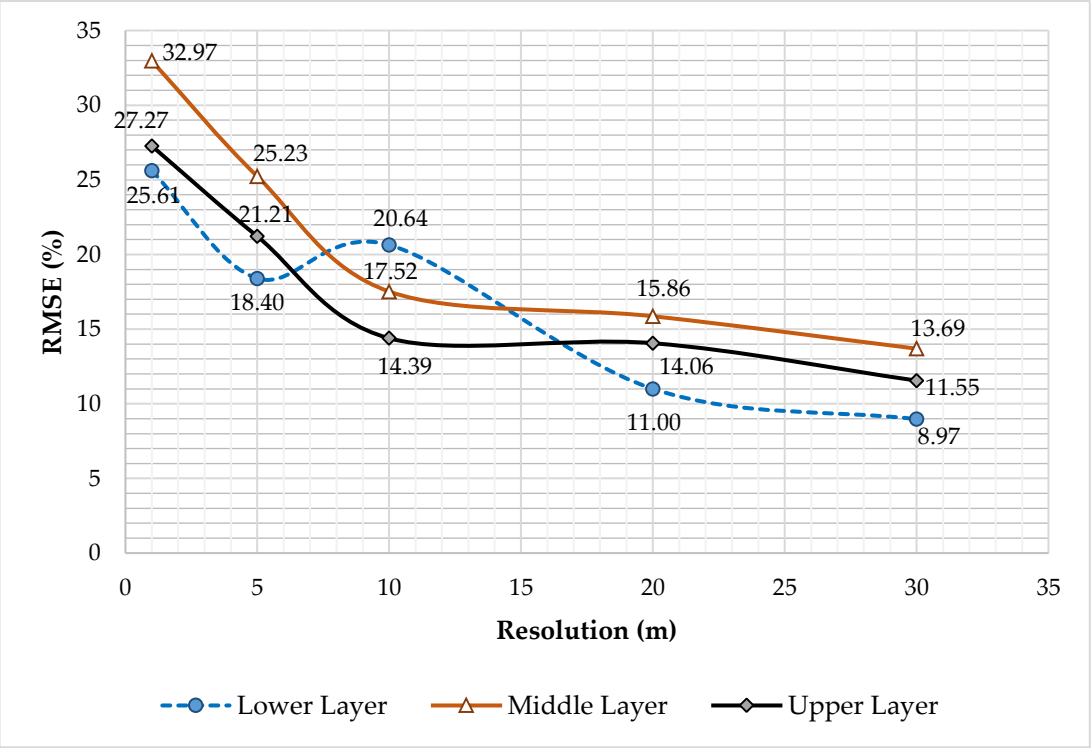


Figure 4.15 Model validation accuracies using the independent validation datasets.

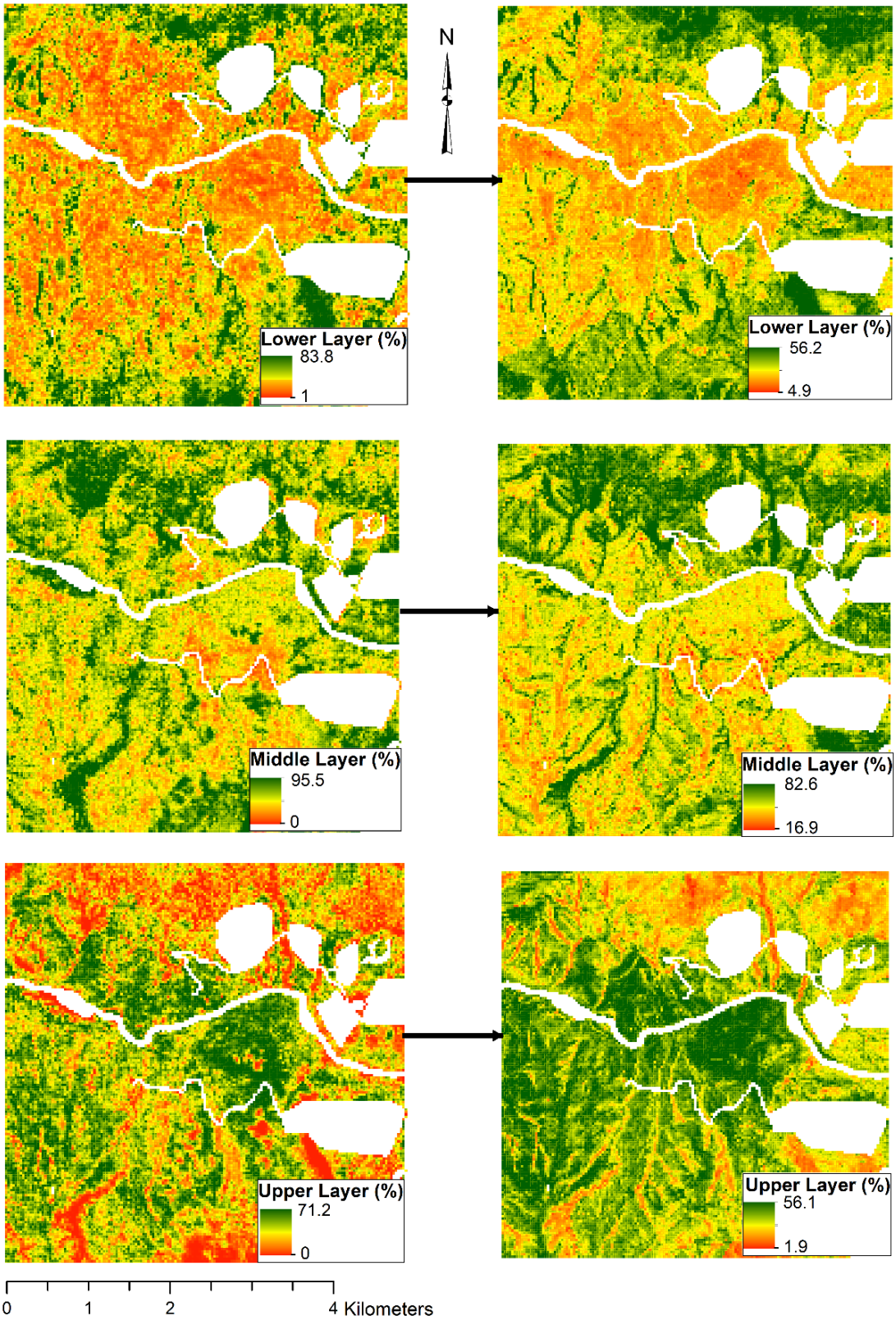


Figure 4.16 Maps showing the density of forest understory layers that were observed (left) and predicted (right) for a 4.9 km by 4.9 km study area using 30 m resolution. Roads, rivers and harvested forests have been masked out and appear as white.

4.4 Discussion

This study deployed LiDAR-derived topographic attributes, and a geology map for modelling and mapping three understory density layers using five spatial resolutions in mature wet eucalypt forests. Since high-density LiDAR data acquisition is extremely expensive, prediction of forest structure based on attributes derived from a DTM and geology map could be very useful for informing forest management. In the following sub-sections, this study first discusses how RF models performed for the three understory layers and compares the predicted and validated models of different pixel sizes. Secondly, the effects of key topographic attributes on the understory structural layers are evaluated. The impacts of different spatial resolutions on the variable importance of RF modelling are discussed, and finally, the implications of this study are presented.

4.4.1 Model performance and validation

Distinguishing the heights of a wet eucalypt forest study site into lower, middle and upper layers for further analyses (Wing et al. 2012), thirteen predictor variables were used to predict those three understory density layers at five different resolutions (i.e., 1 m, 5 m, 10 m, 20 m, and 30 m). Wood et al. (2011) reported that it was important to consider the optimum resolution or pixel size of satellite data to explore the effects of topographic attributes. This study could not use field plot data for training and validating RF models, and instead, the high-density airborne LiDAR data were taken as a proxy for understory vegetation layers. Although a DTM derived through photogrammetric techniques was available for the study site, I attempted to test whether DTMs from airborne LiDAR data would be a good source of information for predicting canopy densities of wet eucalypt forest. This approach aimed to establish a relationship of understory layers with topographic attributes and geology data. The 30 m resolution dataset had the highest prediction and validation accuracy and so was considered the best model for modelling and mapping the structural layers using topographic attributes in a wet eucalypt forest. The models for the upper layer had the highest prediction accuracies with the MRMSE value of 7.61%, followed by the lower layer (MRMSE = 8.76%) and the middle layer produced the lowest predictive power (MRMSE = 9.76%). In line with these results, Latifi et al. (2016) also predicted overstory and understory better than the midstory. Compared to the 30 m resolution, the 1 m resolution dataset had the lowest prediction accuracies with MRMSE values more than double those for 30 m. Most of the trees in the study area possessed a crown diameter

of around 30 m that matched the pixel size of 30 m dataset. The 30 m resolution also contains more canopy elements in each pixel compared to the 7.5 m or 1.6 m resolution. Therefore, the 30 m resolution provided the best result. In line with this result, Zald et al. (2014) showed that LiDAR metrics characterized trees with reduced accuracy as the plot size decreased. Azaele et al. (2012) found that objects (e.g. tree crowns) that occupied a pixel were measured at higher accuracy. Thus, this result matched the results of previous studies.

The validation dataset for the preferred 30 m resolution produced the highest model accuracy for the lower layer (MRMSE = 8.97%) followed by the upper layer model (MRMSE = 11.55%), and the middle layer had an MRMSE value of 13.69%, i.e. the lowest validation accuracy. In contrast to the validation accuracies, most previous studies found that the upper layers had higher accuracy than the middle and lower layers. The reasons behind the lower prediction of the middle layer should be that hyperspectral image data primarily contain the spectral signal of the dominant overstory. LiDAR pulses that mostly hit the upper layers and lower layers with small trees, shrub and herbs, but harder to penetrate their pulses in the middle layer of the canopies resulting in the lower accuracy. Spatial variability in seed availability, forest age, canopy closure, and light environment are likely to be important drivers for the successful development of mid-story species. These variables are unlikely to be captured in terrain derivatives and a geological map, and therefore resulted in a relatively modest RF model performance. In the previous studies, for example, Morsdorf et al. (2010) also achieved overall accuracy of 80% to 90% for dominant layers and around 48% for sub-dominant layers, and they mentioned that accuracies could be lower in more complex plots. Wing et al. (2012) found understory vegetation cover with the accuracies ranging from 20% to 45% and projected the accuracies of 77% by combining leaf-off and leaf-on datasets in their deciduous forest ecosystem. Also, Suchar and Crookston (2010) reported adjusted R^2 values of 0.22 and 0.24 for the percent shrub cover models for the forest ecosystems of north-western United States, and they argued that shrub-herb cover is more heterogeneous than overstory cover attributes. Thus, the models in this study performed well compared to those from other forest types and regions. In the study site, varieties of ferns, mostly *Dicksonia antarctica* (soft tree fern or manfern) frequently covered the lower layer of the forests, whereas the tall wet *Eucalyptus* species, mainly *Eucalyptus obliqua* (messmate, stringybark or Tasmanian oak) are dominant trees. Thus, LiDAR-derived topographic attributes were capable and useful variables for predicting the lower layer of understory structure in a wet eucalypt forest ecosystem.

4.4.2 Influences of topographic attributes on understory structural layers

Recognizing important topographic attributes to control vegetation distribution is vital for sustainable forest management (Wang et al. 2015). This study identified geology, solar radiation, topographic position index, aspect, and SAGA wetness index as having the best descriptive power to further explore the effects of topographic attributes on the understory structural layers.

Soil nutrients are one of the major factors that help to grow highly heterogeneous vegetation communities (Ridolfi et al. 2008) and are important drivers of understory vegetation species composition and growth (Kasel et al. 2017). The availability of soil nutrients is strongly determined by geology type. Geology is the major driver of forest structure and stratification and its impact is well known. Even specific types of geology, such as Jurassic dolerite, have higher influence on the understory layers of the forest structure, clearly shown by partial dependence plots on p.73 Section 4.3.4. The Geological data layer appears to act as a stratification layer in the RF model that stratifies the landscape for which the fine-scale variability is then predicted by other variables, such as the terrain derivatives. Kirkpatrick et al. (2014) investigated the influences of vegetation, topography, climate and parent materials on the characteristics of alpine soils in Tasmania and New South Wales in Australia and found parent material was highly influential to model soil attributes. Simonson et al. (2014b) predicted shrub understories could be more developed because of increasing soil nutrients and water availability. Permian sediments (P) demonstrated the highest influence on the upper layer vegetation composition and did not show a significant impact on the lower layer and middle layer. By contrast, Jurassic dolerite (Jd) had the highest influence on the lower layer and middle layer vegetation structure and very low influence on the upper layer. So, forest with Jurassic dolerite (Jd) had a greater influence on the lower and middle density layers, whereas that had a lower influence on the upper density layer of the model. It should be noted that within the small study region, Jurassic dolerite was confined to high elevation sites, so elevation may have been confounded with geology in this case. This would need to be confirmed by research in a broader study area that included dolerite at low elevations. Dolomite (Lwcdss) and Permian (P) had a greater influence on the upper density layer and lower influence on the middle and lower density layers of the models indicating well developed forest. To investigate specific geology types and their relationships with lower layer vegetation structure is beyond the scope of this research, although this type of research could be an interesting topic for future research. It is relevant to consider

soils for forest management and restoration (Wang et al. 2015). Soils with fewer nutrients theoretically support reduced lower layer growth compared to soils with higher nutrient levels. Cambrian silicious sediments (Cals) and dolomite (Lwcdss) would be the most nutrient poor substrates, followed by Permian sediments while limestone would be best.

Solar radiation was considered the second influential attribute in determining the vegetation structure. Like geology, solar radiation had varying influence on the density of the lower layer, middle layer, and upper layer vegetation. Jenkins and Coops (2011) also pointed out the importance of higher insolation levels for canopy characteristics derived from LiDAR in eucalypt forest in New South Wales, Australia. The influence of solar radiation is most likely to be indirect. Areas of high solar radiation are likely to be warmer and drier than areas with low solar radiation. This is likely to increase the frequency and intensity of fires, and potentially may affect vegetation through differences in water demand in summer. The more open the upper vegetation layers, the more light that penetrates to lower layers, which determines the positive associations with understory species richness (Weisberg et al. 2003). Simonson et al. (2014b) found less developed plant canopies due to reduced light transmission. They found canopy cover and density were higher on north-facing aspects receiving less solar radiation, whereas canopy density was negatively associated with topographic position index. In contrast, in this study, partial dependence plots for topographic position index did not demonstrate a significant impact on understory vegetation models. However, Simonson et al. (2014b) showed how *Quercus canariensis* was favoured in the more mesic conditions of gullies and north-facing slopes. This attribute negatively influenced canopy heights of trees in the south-facing plots.

In general, the partial dependence plots demonstrated that the lower the value of the SAGA wetness index (SWI), the greater its impact on understory density layers for the lower- and middle-layer conditions. Although the contribution of SWI was not clear in this study, the wetness index helped to characterize the biological distribution and species diversity elsewhere (Moore et al. 1988). In this study, geology, solar radiation, topographic position index, aspect and SAGA wetness index were crucial components for understory structural development, however, neither of these components affected all their relative influence on the three vegetation layers varied. Wang et al. (2015) found that slope, aspect and surface curvature were important topographic attributes that affected vegetation distribution. Thus, all these topographic attributes played

important roles in explaining vegetation dynamics for particular locations, contributing to ecological studies and sustainable management of forests as a whole.

4.4.3 Impacts of spatial resolution on the variable importance

Spatial resolutions of topographic attributes derived from LiDAR data play significant roles in predicting understory structural layers of wet eucalypt forest. This study assessed the utility of resolutions of 1 m, 5 m, 10 m, 20 m, and 30 m. Previous researchers, for example, Deng et al. (2007) used spatial resolutions ranging from 20 m to 480 m by resampling a DTM to calculate six topographic attributes for classifying landforms and displayed different patterns among various landform classes. Kienzle (2004) extracted the resolutions ranging from 5 to 100 m from a DEM to find out their effects on topographic derivatives and explored that slope and elevation showed a positive relationship, but all other topographic derivatives could not be represented convincingly from a coarser DEM. The value of slope varies with spatial resolution, which would, therefore, have a big impact on the topographic analysis.

In this study, understory models yielded different findings for different resolutions. This may relate to the topographical complexity of the Warra Supersite forest where wildfire disturbance is a major driver of vegetation patterning (Hickey et al. 1999). The importance of topographic attributes not only changed due to changing the resolutions but also due to different understory layers, for example, TPI had the highest importance for the upper layer, but lower importance for the middle and lower layers in 1 m resolution. Thus, all the predictor variables produced different importance scores with different canopy layers and at different resolutions. With this regard, Nijland et al. (2014) reported that the selection of variables changed when LiDAR-derived canopy layers and landscape attributes were added to the models.

4.4.4 Management implications

Many temperate forests are susceptible to natural disturbances like wildfires and are actively managed for timber production, yet long-term effects of disturbance and forest management on understory plant communities are to some degree overlooked (Thomas et al. 1999). A new approach, i.e. linking understory vegetation structure with topographic attributes and a geology dataset for primary forests, is relevant for sustainable forest management, fire hazard prediction and biodiversity conservation in wet eucalypt forest ecosystems. This approach for predicting understory vegetation structure avoids the use of expensive and time-consuming field inventory data. However, this technique still has limitations for practical use, because plant species

present at the specific sites cannot be predicted without the detailed knowledge of the upper layer affecting species composition and understory vegetation (Thomas et al. 1999). Furthermore, this study was restricted to mature forests, and may not be relevant to younger successional ages, including forests that were regenerated following timber harvesting. Nevertheless, there are several studies highlighting the usefulness of LiDAR data to predict habitat suitability of forest dwelling species (Morsdorf et al. 2010), bird species richness (Goetz et al. 2007), and assessment of biodiversity (Ferris and Humphrey 1999). Thus, this approach could also be applied for fuel-load management and fire behaviour modelling using the appropriate topographic variables.

4.5 Conclusions

The results of this study demonstrated that LiDAR-derived topographic attributes and geology information could be used to predict the understory density of three canopy layers for a wet eucalypt forest. This study has shown the significance of topographic attributes and geology on the understory canopy layers of wet eucalypt forest. This approach could be applied remotely to landscapes where a DTM is available, but high-density LiDAR data and field data on vegetation are not available. The model presented in this study utilised a number of spatial resolutions for different understory density layers.

Based on this study, the following conclusions were made:

1. Variable importance changes with changing resolutions of datasets as well as the understory canopy layers. In general, geology performed the best for predicting vegetation density of the lower and upper layers while solar radiation, followed by geology was best for predicting the middle layer in a wet eucalypt forest. Plan and profile curvature produced had less influence on model performance, whereas topographic position index had moderate importance.
2. This research examined *mtry* and *ntree* rigorously and found *mtry* of 13 with *ntree* of 500 provided the lowest error.
3. Based on model accuracies, the 30 m resolution dataset was best for understory structure modelling using topographic attributes and a geology map for a wet eucalypt forest landscape. However, the optimum resolution would need to be determined separately for other landscapes, ecological processes, and habitat types, i.e. agriculture, hydrological studies, etc.

4. Using selected topographic attributes and the mapped geology, the upper layer provided the best model fit with MRMSE values of 7.61% (R^2 = 0.82), and the middle layer structure had the lowest predictive power with the MRMSE value of 9.76% (R^2 = 0.77). The validated models for the lower layer provided the best model fit with an RMSE value of 8.97%, followed by the upper layer (RMSE = 11.55%) and the middle layer had the lowest model performance with RMSE of 13.69%.
5. It was observed that the same predictor variable affected the model in different ways (as shown in the partial dependence plots) depending on the understory structural layers and spatial resolutions used for modelling purposes.

This study has also highlighted useful topics for future research:

1. This approach needs to be applied over larger geographical areas with more diversity of topographic features to widen the scope of the present study (Ziadat 2005).
2. The model presented in this study depended on high-resolution airborne LiDAR data that could be costly and so non-repeatable. Therefore, applications of lower-resolution 'operational' LiDAR and multispectral remote sensing data could be deployed to model understory density layers (the research topic of Chapter-5).

Overall, these results could be used to predict understory vegetation structural layers and biomass, assess forest health and biodiversity and determine fuel loads. This study should be useful for foresters and ecologists and has the potential to contribute to sustainable forest management and biodiversity conservation.

Chapter 5

Assessing the robustness of multispectral satellite imagery with LiDAR topographic attributes to predict vertical structure across multiple scales in a wet eucalypt forest

Abstract

The vertical arrangement of forest canopies is a key component for making forest management decisions for timber production, fire management, and wildlife habitat. Recent advancements in remote sensing technologies facilitate modelling and mapping of forest areas from fine scales to large geographical areas, especially for inaccessible and complex forests. This study aimed to assess the robustness of multispectral satellite imagery (i.e. WorldView-3 and Landsat-8 Operational Land Imager (OLI)) and topographic attributes derived from airborne LiDAR data, downsampled to operational point density of 4.94 points/m^2, to predict the density of three canopy layers. This was conducted for three pixel sizes in a wet eucalypt forest located at the Warra Supersite in Tasmania, Australia using random forest regression modelling. This research applied four steps: (a) data pre-processing, (b) extraction and selection of response and predictor variables, (c) random forest regression and variable importance assessment, and (d) assessment of model performance and validation. Using spectral bands, spectral indices, texture features, and topographic attributes as predictor variables, we compared and examined the predictive power of the thirteen data schemes at three different pixel sizes (1.60 m, 7.5 m and 30 m). The result of this study demonstrated that the schemes of the 30 m Landsat-8 (OLI) dataset provided better overall model accuracy than the WorldView-3 dataset across all three pixel sizes (R^2 values from 0.15 to 0.65). Our results showed that the model accuracies increased with an increase in the number of predictor variables. The models validated using independent datasets also confirmed the robustness and transferability of the developed models, as the differences between the predicted accuracies and validated accuracies of the models were less than 5.67%. However, in situations where the lower layer of a wet eucalypt forest is of particular interest to forest managers and researchers rather than the middle and upper layers, the WorldView-3 could be more appropriate than the Landsat-8 (OLI) dataset. Moreover, the topographic attributes derived from the simulated operational LiDAR data contributed significantly to the modelling schemes of 30 m pixel size datasets. WorldView-3 datasets with the pixel size of 1.6 m were not able to predict the vertical forest structure as the R^2 values ranged from 0.0 to 0.08. Overall, the spectral indices and texture features derived from the freely available and analysis-ready Landsat data products integrated with the simulated operational LiDAR data can provide valuable information on forest structure for sustainable management and monitoring of wet eucalypt forest.

Keywords: Airborne LiDAR; Multispectral satellite imagery; Vertical forest structure; Spectral indices; Texture features; Topographic attributes; Random forest

5.1 Introduction

Vertical forest structure is an important component for making forest management decisions affecting microclimate, nutrient cycles, vertebrate animal and arthropod behaviour (Barkman 1988; Zehm et al. 2003), the habitat requirements of different organisms (Rutten et al. 2015), and the distribution of fuels and fire behaviour (Brokaw and Lent 1999; Campbell et al. 2018; Ehle and Baker 2003). The vertical forest vegetation layers are usually used to represent a *forest stand structure* and can be maintained or altered by managing forests or by wildfire (Brokaw and Lent 1999). Detailed knowledge of vertical stand structure is required to manage forests for multiple purposes, such as wildlife habitat, timber production and fire management (Hudak et al. 2006), and to understand important ecosystem functions such as hydrological cycling and carbon sequestration (Dash et al. 2016). Also, the overstory composition and structure influence understory vegetation due to the modifications of sun light, soil moisture and nutrients (Barbier et al. 2008; Bartels and Chen 2013; Campbell et al. 2018), and the relationship among canopy layers is complex (Coll et al. 2011). However, measuring vertical forest structure directly in the field is a challenge for forest managers (Wilkes et al. 2015) and is expensive, time-consuming, and even difficult in areas that are inaccessible (Clawges et al. 2008). Thus, ecologically meaningful and robust measures of vertical forest structure for remote, inaccessible and large geographical areas have been lacking (Bergen et al. 2009; Culbert et al. 2013; Zimble et al. 2003).

Remote Sensing (RS) satellite data facilitate modelling, mapping and understanding of ecosystems (Lefsky et al. 2002) and are an inexpensive technique for modelling and mapping forest structural attributes in large geographical areas (Corona et al. 2002; Masek et al. 2015; Maselli et al. 2005; Ozdemir and Karnieli 2011). Mapping forest structure attributes using remotely sensed data has been of interest for over four decades (Kayitakire et al. 2006; Masek et al. 2015) and provides a variety of information to forest managers, including impacts of disturbance and indicators of biodiversity (Pasher and King 2010). Although aerial photographs were previously used for mapping forest stands, those techniques were highly subjective, time-consuming and depend on the experience of the interpreter (Bolton et al. 2018; Kayitakire et al. 2006; Schroeder et al. 2007). With the advancements of technologies, an automated image analysis technique was explored to retrieve forest stand structural attributes by Kayitakire et al. (2006), and satellite imagery was found to meet the needs of spatially complete information about forests over large geographical areas (Zald et al. 2014; Zald

et al. 2016). Thus, moderate resolution, e.g. Landsat data, plays a crucial role in mapping the species composition of forests (Masek et al. 2015) and the Landsat-8 Operational Land Imager (OLI) pushbroom data contains rich information (Dube and Mutanga 2015). Similarly, commercial high-resolution satellite data, e.g. WorldView-3, has raised great interest among forest managers and planners for potential application for forest inventory purposes (Kayitakire et al. 2006). Although traditional passive remote sensing techniques have demonstrated the potential to provide useful information on forest structure attributes based on spectral reflectance and derived vegetation indices (Cohen and Goward 2004; Gebreslasie et al. 2010; Schroeder et al. 2007), these techniques are not currently sufficient to describe vertical three-dimensional (3D) forest structure for forestry and ecological applications.

As an alternative to passive remote sensing techniques, Light detection and ranging (LiDAR), also called airborne laser scanning (ALS), is an active remote sensing technology that can directly measure the 3D vegetation structure characteristics by penetrating canopy layers. This provides a valuable tool in forest inventory and other ecological applications (Bigdeli et al. 2018; Dash et al. 2016; Lefsky et al. 2002; Wilkes et al. 2016), for example, assessment of structure response to a range of bushfire events (Bolton et al. 2015), wildlife-habitat relationships (Clawges et al. 2008; Vierling et al. 2008) and habitat quality (Hinsley et al. 2006), and bird species richness (Lesak et al. 2011). Moreover, LiDAR sensors are sensitive to vertical forest structure even in the case of high biomass forests whereas the optical multispectral satellite data lose their sensitivity and saturate (Hudak et al. 2006; Hudak et al. 2008; Masek et al. 2015), and cannot provide information for underneath the canopy cover (Arroyo et al. 2010). However, the cost, affordable pulse density and complex processing of LiDAR point clouds have limited operational applications (Hudak et al. 2008). Nowadays, there is considerable interest in the integration of LiDAR and multispectral satellite data by land management agencies, forest industries, and public groups to retrieve forest structural attributes (Cohen and Spies 1992; Dash et al. 2016; Wilkes et al. 2015). Considering the different aspects of vertical forest structure, no single remote sensor can exhibit all the required information relevant to forest managers, and the integration of multispectral satellite imagery and LiDAR data could advance prediction and mapping of forest structure characteristics (Hudak et al. 2006; Popescu and Wynne 2004). Most of the developed models related to forest structure have not been transferred from research sites to test their robustness for broad-scale application (Sumnall et al. 2016).

Vertical forest structure is scale dependent, therefore understanding the role of spatial scale is essential to transfer the information from fine scales to broad scales, and vice versa (Levin 1992; Wilkes et al. 2016). The successful application of remote sensing data depends on the optimum scale or an appropriate spatial resolution (Kamal et al. 2014). Scaling also supports the monitoring of multiscale biodiversity attributes (Kunin et al. 2018). Upscaling (also called data aggregation) is typically required to fill the scale gap between remote sensing satellite measurement and field requirements for large-scale models (Hong et al. 2009). Previous studies have highlighted that aggregation to a coarser resolution exposes a particular spatial model (Zhang and Montgomery 1994), and a decrease in spatial resolution loses valuable information (Carmel et al. 2001). The local scale must be scaled up to the entire landscape in modelling. Thus, the assessment of optimum spatial resolution (pixel size) remains an unsolved research issue, especially for heterogeneous ecosystems and is challenging for ecologists (Zhang 2007). Multi-scale remote sensing, therefore, has enormous potential to resolve ecological issues of scales (Chave 2013).

Spectral, contextual and textural features are three fundamental components typically used to interpret an image. Spectral features highlight the average tonal variation in various bands, and textural features possess information about the spatial distribution of the tonal variations (Haralick et al. 1973). Image texture is an important source of information for different aspects of forest stand structure, e.g. tree density (Dube and Mutanga 2015; Wood et al. 2012). A number of previous studies have demonstrated the use of texture features derived from high resolution multispectral satellite data (e.g. WorldView-2) (e.g. Dube and Mutanga 2016; Eckert 2012) and medium resolution multispectral sensors (e.g. Landsat products) (Dube and Mutanga 2015; Phua et al. 2017; Wood et al. 2012; Xu et al. 2011) to estimate various forest structural characteristics. Similarly, spectral vegetation indices relying on spectral bands, mostly the red and near-infrared bands, are well-known to enhance the sensitivity of vegetation properties (e.g. Clevers 2014; Glenn et al. 2008). Vegetation indices are simple and are the oldest and most widely used for the estimation of biophysical and biochemical variables (Verrelst et al. 2015).

Topography plays a significant role in the development of vertical forest structure affecting microclimate and edaphic factors (Gracia et al. 2007). Topography attributes could be distinguished into primary and secondary (or compound) attributes. Primary attributes can be directly derived from a digital terrain model (DTM) (e.g., elevation, slope, plan and profile curvature) and secondary attributes can be calculated indirectly

involving the combinations of primary attributes with other spatial variables (Moore et al. 1993). The studies of topographic attributes provide a basis to understand ecological relationships with vegetation, landform, and soils (Odom and Henry McNab 2000), whereas assessment of topographic effects on the different layers of forest structure has rarely been attempted (Simonson et al. 2014b). Topographic attributes often have the greatest impact on the per-pixel classification of image data, but vary with image spectral and spatial characteristics (Cohen and Spies 1992). In addition, geology or soil maps are valuable tools for natural resource management and may be informative for predicting vegetation structure (Moore et al. 1993).

Random forest (RF) is a non-parametric machine learning technique that has been widely used in forest inventory, planning, and ecological research. The RF model samples the training data randomly and iteratively grows a large group of decision trees called a 'forest' for both Classification And Regression Trees (CART). This property of the RF model distinguishes it from other techniques. This algorithm subsets the predictor variables and prevents problems with overfitting, and also derives variable importance metrics (Breiman 2001). Several studies have used RF models, as for example, Zahedi et al. (2017) compared RF, ANN (Artificial Neural Network) and Support Vector Machine (SVM) and predicted soil depth using environmental variables derived from a DTM, and found the RF model was superior to the other two modelling approaches. Latifi (2012) found RF to be superior to other imputation methods such as Most Similar Neighbourhood (MSN), Euclidean distance and Mahalanobis distance. Hudak et al. (2008) compared Canonical Correlation Analysis (also known as Most Similar Neighbour or MSN), Canonical Correspondence Analysis (also known as Gradient Nearest Neighbour or GNN), Independent Component Analysis (ICA), Mahalanobis distance, normalized and unnormalized Euclidean distance and RF, and concluded that RF was the most robust and flexible among all the mentioned techniques.

Previous studies have also demonstrated that the integration of LiDAR and multispectral satellite data can sometimes provide higher accuracies than either product in isolation, although this is not always the case. For example, Zald et al. (2016) predicted forest structure and aboveground biomass combining LiDAR-derived metrics and Landsat products and characterised key ecological patterns with random forest and nearest neighbourhood imputation mapping. Dash et al. (2016) characterised forest structure combining LiDAR, RapidEye, and environmental variables, and concluded that LiDAR metrics provided the most useful stand attributes, and the

combination with RapidEye provided negligible improvements compared with the LiDAR data alone. Wilkes et al. (2015) studied upscaling approaches using LiDAR and Landsat thematic Mapper (TM) satellite image-derived textural and vegetation indices in a heterogeneous forest in Victoria, Australia. No improvements were achieved in canopy height estimation in their upscaling approach using the RF model. Bolton et al. (2015) conducted their research in Canada's western boreal forests integrating airborne LiDAR and Landsat Time Series (LTS) data to assess the structural complexity of burned stands using measures of canopy roughness and the distribution of LiDAR returns that provided the evidence of young, even-aged structure after a new overstory was created. Zald et al. (2014) investigated in the central Oregon Cascades, USA combining LiDAR-derived topography and vegetation metrics with tasseled-cap indices and disturbance history metrics derived from LTS data as predictor variables for the structural studies of forest vegetation. They found that the LTS and LiDAR indices could not improve the prediction of forest structure and species respectively and suggested the need to prioritize the forest vegetation attributes for a specific application. Erdody and Moskal (2010) used LiDAR and high-resolution colour near-infrared aerial imagery for predicting canopy fuel metrics in a fire-prone ponderosa pine dominated forest in Washington State and showed that the fusion of sensors can improve the accuracy of canopy metrics. They suggested that the LiDAR was superior to near-infrared imagery because LiDAR represented the vertical forest structure.

There is relatively limited research that has compared the model accuracies predicting vertical forest structure generated by combining LiDAR-derived topographic attributes with other sources of remote sensing satellite data, especially in complex mixed forests (Dash et al. 2016; Johansen et al. 2010; McInerney et al. 2010; Mikita et al. 2013). This is an important topic of research to extend the scope of LiDAR-derived variables while keeping the acquisition costs low for large geographical areas (McInerney et al. 2010). Also, no single sensor is perfectly suited for the estimation and mapping of forest characteristics (McInerney et al. 2010). Indeed, measurements of direct field inventory collect highly detailed forest measures and ecological data, but they are limited to sampled locations and lack a complete spatial coverage for mapping forest structure often required for forest managers and planners (Zald et al. 2014). Furthermore, research is also required to perform the retrieval of topographical changes and their effects on heterogeneous stands (Kayitakire et al. 2006), and comparison of different texture measures and scales (Ferro and Warner 2002). Although the relationships between image texture metrics and forest structure attributes can be used to characterise complex forest structure and enhance the fine vegetation properties, the

uses of vegetation indices particularly in the case of closed canopies are difficult and challenging (Dube and Mutanga 2015). In addition, the problem of scaling up for the prediction of larger areas is one of the central problems in ecology (Cipriotti et al. 2016). Keeping these research gaps in mind, the primary objective of this study is to assess the robustness of multispectral satellite imagery with or without topographic attributes derived from a DTM to predict vertical structure across multiple scales in a wet eucalypt forest located at the Warra Supersite in Tasmania, Australia using random forest regression modelling. Moreover, this study assessed whether the selected vegetation indices and texture features improved the prediction accuracy of the vertical structure attributes, i.e. densities in three canopy layers used as response variables. Multispectral satellite datasets (i.e. WorldView-3 and Landsat-8, OLI) were processed and analysed to retrieve the values of spectral bands, vegetation indices, and texture features separately and fused these with the topographic attributes derived from a DTM and ancillary data (i.e. a geology vector layer). To provide insights into the different sources of datasets owing to different scales, this research investigates how model outcomes are influenced by three different pixel sizes.

5.2 Methods

5.2.1 Study area and simulated datasets

The details of the study area and datasets were presented in Chapter 2. This study used cloud-free WorldView-3 satellite multispectral imagery, Landsat-8 (OLI) land surface reflectance imagery and discrete return airborne LiDAR point clouds. In this study, the discrete return airborne LiDAR data of 28.66 points/m^2 were thinned to reduce the point density to 4.94 points/m^2 and spacing of 0.45 (all returns) to simulate operational low-density LiDAR data typically available for forest surveys. Although direct forest inventory provides detailed information on forest vertical structure, this study did not consider field plot data, and used an approach solely depending on remote sensing data, simulating operational LiDAR data from a high-density LiDAR data for predicting vertical forest structure. WorldView-3 satellite imagery were resampled from 1.60 m for the visible near-infrared (VNIR) 8 bands to 7.5 m using pixel aggregation/averaging (e.g., Ferro and Warner 2002) in ENVI software version 5.3.1 to match the spatial resolution of the SWIR 8 bands of the same WorldView-3 imagery and then used all 16 bands for further analyses. Similarly, all the 16 bands of WorldView-3 imagery were again resampled to 30 m pixels to match the spatial resolution of the Landsat 30 m (OLI) dataset. The pixel aggregation is an appropriate

method to process remote sensing imagery because a pixel value is averaged over the associated area of the coarser resolution pixel as a measure of central tendency of all the pixels contained within the spatial grid cells (Blan and Butler 1999; Colin et al. 2018). The WorldView-3 imagery was resampled to examine the performance of the resampled datasets and to scale up the vertical forest structure models with the larger pixel sizes. Thus, this study focused on the fusion of multispectral satellite imagery and DTM derivatives to predict vertical forest structure.

5.2.2 General description of the method

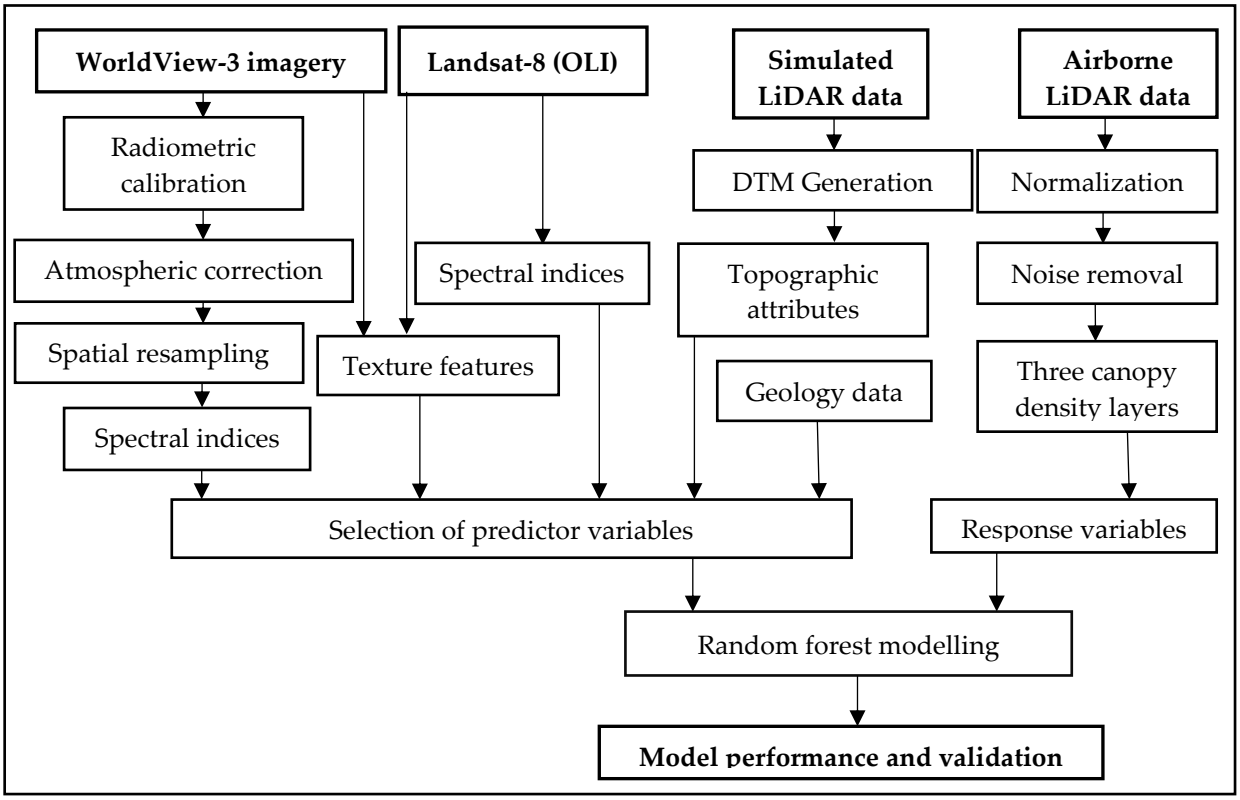


Figure 5.1 Workflow diagram of the study.

The workflow of this chapter is presented in Figure 5.1 and the detailed technical aspects are described in subsequent sections and sub-sections. The whole approach was separated into four steps: (a) data pre-processing, (b) extraction and selection of appropriate predictor variables, (c) random forest modelling, and (d) assessment of model performance and validation. The LiDAR data pre-processing included noise removal, derivation of three DTMs (1.60 m, 7.5 m, and 30 m), normalization, and extraction of vegetation densities in three canopy layers for each of the 1.60 m, 7.5 m and 30 m spatial resolutions. WorldView-3 multispectral satellite imagery required radiometric calibration, atmospheric correction, and extraction of spectral indices and eight texture features from each band. Similarly, the spectral indices and eight texture

features were extracted from each band of Landsat-8 (OLI) land surface reflectance imagery. The selected predictor variables were then deployed for non-parametric random forest modelling. Finally, model performance was assessed and validated.

5.2.3 Data pre-processing

To examine the effects of scales on the prediction accuracies, this study was conducted on the varying spatial resolutions combining the derivatives of LiDAR and multispectral satellite imagery, i.e. WorldView-3 and Landsat-8 datasets. LiDAR point clouds were processed using LAStools (Academic version 180907). The study area was first merged and clipped from the original 100 tiles. The point clouds were classified into the ground and non-ground categories using *lasground* and then normalized using *lasheight*. The normalized data was filtered to provide noise-free data for further analysis (Onojeghuo and Onojeghuo 2017). Finally, the three height density rasters were derived using *lascanopy* to use them as response variables. The three height density rasters of ≥2 to ≤ 10 m, > 10 m to ≤ 30 m and > 30 m to ≤ 50 m were derived using high-density discrete return LiDAR point clouds (approximately 30 points/m^2). As mentioned in Chapter 4, the vertical structure of the forest area was classified into three canopy layers: a lower layer (≥ 2 to ≤ 10 m), a middle layer (> 10 m to ≤ 30 m) and an upper layer (> 30 m to ≤ 50 m), based on expert knowledge of the vegetation communities. The lower layer mostly contains tree ferns and low understory shrubs and trees of up to 10 m in heights, the middle layer contains mixed tree species, and the upper layer consists of dominant trees, primarily eucalyptus overstory. Next, the original LiDAR point clouds were thinned to simulate the operational LiDAR data (approximately 5 points/m^2), and the three DTMs (1.60 m, 7.50 m and 30 m) were generated using *blast2dem*. The pulse returns below 2 m above the ground layer were discarded so that only representative canopy hits were considered, avoiding low ferns, small shrubs and coarse woody debris (e.g., Bolton et al. 2015; Ehlers et al. 2018; Hudak et al. 2008; Jaskierniak et al. 2011; Johansen et al. 2010; Latifi et al. 2012b; Packalén and Maltamo 2006; Wilkes et al. 2016).

Like other multispectral remote sensing data, the raw WorldView-3 satellite data needs to be converted from digital number (DN) values to top-of-atmosphere spectral radiance and then to surface reflectance before performing further analysis (Kuester 2016). The correction of atmospheric effects is important (Song et al. 2001), so an atmospheric correction was performed. But, because the gain, offset and irradiance values were not found in the .IMD metadata for WorldView-3 data provided by the

vendor, radiometric calibration was performed using the information reported in Kuester (2016) (Table 5.1).

Table 5.1 Absolute radiometric calibration adjustment factors and irradiance values for WorldView-3 as of 1/29/2016.

Band	**Gain value**	**Offset value**	**Solar irradiance value (W-M^{-2}−μm^{-1}) (Thuillier et al. 2003)**
Coastal	0.863	-7.154	1757.89
Blue	0.905	-4.189	2004.61
Green	0.907	-3.287	1830.18
Yellow	0.938	-1.816	1712.07
Red	0.945	-1.350	1535.33
Red-Edge	0.980	-2.617	1348.08
NIR 1	0.982	-3.752	1055.94
NIR 2	0.954	-1.507	858.77
SWIR 1	1.160	-4.479	479.019
SWIR 2	1.184	-2.248	263.797
SWIR 3	1.173	-1.806	225.283
SWIR 4	1.187	-1.507	197.552
SWIR 5	1.286	-0.622	90.4178
SWIR 6	1.336	-0.605	85.0642
SWIR 7	1.340	-0.423	76.9507
SWIR 8	1.392	-0.302	68.0988

Source: Radiometric use of WorldView-3 imagery prepared by Kuester (2016)

In the process of radiometric calibration, gain and offset values are required that are the absolute radiometric calibration adjustment factors and applied to all pixels of each band of the WorldView-3 imagery to convert raw DNs to top-of-atmosphere spectral radiance (Kuester 2016). The radiometric calibration was generally performed to compensate radiometric errors due to defects in the sensor, scale-angle variations, and system noise so that imagery can represent the true spectral radiance at the sensor (Onojeghuo and Onojeghuo 2017). Atmospheric correction on the imagery was then performed to achieve the top of atmospheric reflectance applying a QUick Atmospheric Correction (QUAC) model. This model performs atmospheric correction on multi- or hyperspectral imagery of all or part of the VNIR and SWIR spectral range ~400 to 2500 nm (Bernstein et al. 2012). The QUAC model determines the compensation parameters directly from the information contained within the scene without ancillary information. WorldView-3 data were re-sampled to 7.5 m to match

the SWIR bands and 30 m to match the Landsat-8 (OLI) dataset. No corrections were performed on Landsat-8 (OLI) surface reflectance imagery except clipping the study area and could be directly used as input to the biophysical models (Flood 2014). Roy et al. (2010) also mentioned that for at-sensor reflectance of OLI imagery, users do not need to perform the non-linear transformation from radiance to reflectance. All the datasets were registered and reprojected to GDA94/MGA zone 55. The registration quality was visually assessed by overlaying a vector coverage map of the study area also done by Kayitakire et al. (2006). As this study focused on predicting vertical forest structure in mature forests, previously harvested sites, roads and rivers were excluded from further analysis.

5.2.4 Response and predictor variables

This study deployed three density layers derived from high-density LiDAR data for the response variables, and thirteen schemes of datasets for the predictor variables combining multispectral satellite imagery and topographic attributes extracted from a DTM. The same twelve topographic attributes and a geology map were used as in Chapter 4: slope, aspect catchment area, and solar radiation, profile curvature, plan curvature, convergence index, terrain ruggedness index, slope length and steepness (LS) factor, SAGA wetness index, topographic position index and stream power index were extracted from the DTMs using SAGA (System for Automated Geoscientific Analyses) GIS software. Refer to Chapter 4 for details.

I derived eight texture features derived from the grey-level co-occurrence matrix (GLCM) (Wood et al. 2012) for each band of the multispectral satellite imagery based on their ability to characterize vegetation structure (Dobrowski et al. 2008; Ge et al. 2006) (Table 5.2). Table 5.3 presents the fifteen spectral indices that were selected. Both the texture features and spectral indices were calculated using the open-source R programming language (R Core Team 2017).

The eight texture features were generated from each band of the Landsat-8 (OLI) data (7 bands multiplied by 8 is 56 image layers). Similarly, the 128 rasters of texture features (16 bands multiplied by 8 texture features) were derived from WorldView-3 data to examine the performance of the texture features to predict vertical forest structural layers. Variable selection is mostly not required in non-parametric modelling (Cutler et al. 2007; Iqbal et al. 2019); however, as this study dealt with a large number of variables (e.g. 100+) dimensionality reduction was carried out (Cadima et al. 2004). In this case, a variance inflation factor (VIF) was applied to remove the collinear

variables using a threshold of 5 (Imdadullah et al. 2016). The VIF identifies collinearity among the predictor variables: the higher the value, the higher the collinearity.

Table 5.2 Selected texture features derived from WorldView-3 and Landsat-8 (OLI) satellite data.

Texture feature	Description	Equations	References
Contrast	Grey level of the two pixels of the same image varies	$\sum_i \sum_j (i-j)^2 \, p(i,j)$	(Kayitakire et al. 2006)
Correlation	Captures how the pairs of pixels are correlated to other pixel pairs	$\frac{\sum_i \sum_j ijp\,(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	(Kayitakire et al. 2006)
Dissimilarity	Two samples vary with the number of grey levels	$\sum_i \sum_j \lvert i - j \rvert \, . \, p(i,j)$	(Soh and Tsatsoulis 1999)
Entropy	Captures the amount of variation in the co-occurrence of the grey level distribution	$-\sum_i \sum_j p(i,j) \log(p(i,j))$	(Haralick et al. 1973)
Homogeneity	measures how close the distribution of elements in the GLCM	$\sum_i \sum_j \frac{1}{1+(i-j)^2} \, . p(i,j)$	(Soh and Tsatsoulis 1999)
Mean	Mean value of intensities over the image	$\sum_{i=2}^{2N} i p_{x+y}(i)$	(Haralick et al. 1973)
Angular second moment	a measure of homogeneity of an image/measures the local uniformity of the grey levels	$\sum_i \sum_j (p(i,j))^2$	(Haralick et al. 1973)
Variance	a measure of "roughness"	$\sum_i \sum_j (i-\mu_i)^2 \, p(i,j)$	(Kayitakire et al. 2006)

Note: Let $p(i,j)$ is the two compared pixels in the image, one with grey level i and the other with grey level j. μ_x, μ_y, σ_x *and* σ_y are the means and standard deviations of p_x and p_y. These eight texture features are widely used in the literature and important variables.

Table 5.3 Selected spectral indices derived from WorldView-3 and Landsat-8 (OLI) satellite data.

SN	Spectral indices	Acronyms	Equations	Reference
1	Green Atmospherically Resistant Index	GARI	$\frac{NIR - [Green - \gamma(Blue - Red)]}{NIR + [Green - \gamma(Blue - Red)]}$	(Gitelson et al. 1996)
2	Green Normalized Difference Vegetation Index	GNDVI	$(NIR - Green)/(NIR + Green)$	(Gitelson and Merzlyak 1998)
3	Infrared Percentage Vegetation Index	IPVI	$\frac{NIR}{NIR + Red}$	(Crippen 1990)
4	Modified Non-Linear Index	MNLI	$\frac{(NIR^2 - RED) * (1 + L)}{NIR^2 - Red + L)}$	(Yang et al. 2008)
5	Modified Soil Adjusted Vegetation Index	MSAVI	$\frac{2 * NIR + \sqrt{(2 * NIR + 1)^2 - 8(NIR - Red)}}{2}$	(Qi et al. 1994)
6	Modified Simple Ratio	MSR	$\frac{\left(\frac{NIR}{Red}\right) - 1}{\left(\frac{NIR}{Red}\right) + 1}$	(Chen 1996)
7	Non-Linear Index	NLI	$\frac{NIR^2 - Red}{NIR^2 + Red}$	(Goel and Qin 1994)
8	Normalized Difference Vegetation Index	NDVI	$\frac{(NIR - Red)}{(NIR + Red)}$	(Rouse et al. 1973)
9	Renormalized Difference Vegetation Index	RDVI	$\frac{(NIR - Red)}{\sqrt{(NIR + Red)}}$	(Roujean and Breon 1995)
10	Optimized Soil Adjusted Vegetation Index	OSAVI	$\frac{(NIR - Red)}{(NIR + Red + 0.16}$	(Rondeaux et al. 1996)
11	Soil-Adjusted Total Vegetation Index	SATVI	$\frac{SWIR1 - Red}{SWIR1 + Red + L} * (1 + L) - \frac{SWIR2}{2}$	(Hagen et al. 2012; Marsett et al. 2006; Torbick et al. 2016)
12	Normalized Burn Ratio (*not for Landsat (OLI) data*	NBR	$\frac{(NIR_{858\text{nm}} - SWIR_{2250\text{nm}})}{(NIR_{858\text{nm}} + SWIR_{2250\text{nm}})}$	(Ji et al. 2011; Key and Benson 2006)

13	Normalized Difference Water Index	NDWI	$\frac{(NIR_{858\text{nm}} - SWIR_{1640\text{nm}})}{(NIR_{858\text{nm}} + SWIR_{1640\text{nm}})}$	(Chen et al. 2005; Ji et al. 2011)
14	Surface Water Capacity Index	SWCI	$(\text{SWIR}_6 - \text{SWIR}_7)/(\text{SWIR}_6 + \text{SWIR}_7)$	(Zhang et al. 2013c)
15	Shortwave Infrared Soil Moisture Index	SIMI	$\sqrt{\frac{SWIR_6^2 + SWIR_7^2}{2}}$	(Zhang et al. 2013c)

5.2.5 Modelling schemes and sampling datasets

This study utilised three density layers derived from high-density LiDAR data to produce training and validation samples. Thirteen combinations of datasets were prepared using topographic attributes, texture features and spectral indices for predictor variables as listed below. To understand the effects of pixel sizes or scales on vertical forest structure modelling using different datasets, all the variables were selected based on their appropriateness and the established applications in forestry related studies (Falkowski et al. 2009; Jaskierniak et al. 2011; Martinuzzi et al. 2009). Ultimately, the random forest models were performed deploying the selected predictor variables on three density layers of vertical forest structure as response variables and compared those models to assess predictive capacity (Singh et al. 2015).

1. Spectral bands (B)
2. Topographic attributes and geology (A+G)
3. Spectral indices (I)
4. Spectral bands, topographic attributes, and geology (B+A+G)
5. Spectral bands and spectral indices (B+I)
6. Topographic attributes, geology and spectral indices (A+G+I)
7. Spectral bands, topographic attributes, geology and spectral indices (B+A+G+I)
8. Texture features (T)
9. Spectral bands and texture features (B+T)
10. Topographic attributes, geology and texture features (A+G+T)
11. Spectral indices and texture features (I+T)
12. Topographic attributes, geology, spectral indices and texture features (A+G+I+T)
13. Spectral bands, topographic attributes, geology spectral indices and texture features (B+A+G+I+T)

In this study, raster datasets were first converted to points and then sampled randomly from all the datasets without replacement in ArcGIS 10.5.1. All the data values covering the harvested sites, rivers, and roads and no data were removed. Out of the remaining point values, 10,000 point locations were randomly selected (Campbell et al. 2018; Criminisi et al. 2011; Wood et al. 2011) and then divided into a model training dataset (50%) and a validation dataset (50%). Again, from the training dataset, 70% of point locations were randomly drawn for model training and 30% for cross-validation (Abdel-Rahman et al. 2013; Kemppinen et al. 2018; Oumar and Mutanga 2014). To achieve robust and stable results and examine the sensitivity of the models, the cross-validation was repeated 100 times (Breiman 1996; Qian et al. 2016; van Galen et al. 2018; Yang et al. 2017b) for predicting vertical forest structure, i.e. densities of three canopy layers. The predicted results were finally validated using the independent validation datasets.

5.2.6 Random forest modelling

Random Forest (RF) regression modelling was used to examine the predictive power of different datasets at different spatial resolutions for vertical forest structure. The RF model is capable of efficiently incorporating a large number of continuous and categorical variables (Wilkes et al. 2015). The RF regression algorithm is a bagging technique, which employs recursive partitioning to divide the input data into many homogenous subsets called regression trees (*ntree*) and then averages the results of all trees. Each tree is independently grown to its maximum size based on bootstrap samples from the training dataset (approximately 67%) without pruning. In each tree, RF selects a random subset of variables (*mtry*) to determine the split at each node (Breiman 2001). In each tree, the ensemble predicts the data that are not in the tree (OOB (out of bag data), approximately 33%)), and by calculating the difference in the mean square errors between the OOB data and dataset used to grow the regression trees. The RF algorithm gives an error of prediction called the OOB error of estimate for each variable (Breiman 2001; Liaw and Wiener 2018; Prasad et al. 2006). The RF algorithm allows us for identification of important predictor variables to predict understory vegetation structure (Martinuzzi et al. 2009). For the RF algorithm, there are two methods of calculating variable importance; one provides a measure of accuracy, called the mean square error (MSE) and the other measure is the Gini index (a measure of node impurity) (Liaw and Wiener 2018). RF calculates OOB errors for model evaluation, thus an independent test dataset may not be required. It can build and evaluate a model from the full dataset and thus take advantage of relatively small

datasets (Freeman et al. 2015). Variable importance is evaluated based on how much worse the prediction would be if the dataset for that variable were permuted randomly (Prasad et al. 2006), and this can be used in feature selection by determining the importance of each variable in the regression process (Freeman et al. 2015). In this study, model performance was tested using the coefficient of determination (R^2).

This study used multiple datasets for different spatial resolutions with the thirteen schemes of predictor variables and three vertical canopy density layers as response variables. The RF modelling analysis processes were conducted in the open source freely available R programming language (R Core Team 2017) using several libraries, including *randomForest* (Liaw and Wiener 2018) and *caret* (Kuhn 2017).

5.3 Results

This section first presents the predicted accuracies of random forest models and then focuses on the patterns for the most successful set of models developed from the 30 m multispectral satellite data (i.e. Landsat-8 and WorlView-3) and the simulated operational LiDAR data. These results compare with those for 7.5 m WorldView-3 (composites of SWIR bands and the simulated VNIR bands) and the 1.6 m WorldView-3 imagery. Finally, the spatial patterns of the predicted forest densities and the validation dataset were compared.

5.3.1 Model accuracy assessment

5.3.1.1 Accuracy for 30 m spatial resolution employing Landsat-8 (OLI) data

The RF model prediction accuracy for the 30 m pixels varied with the inclusion or exclusion of predictor variables, with mean R^2 (hereafter R^2 only) values ranging from 0.15 to 0.65. The R^2 values increased with the increase in the number of predictor variables in most of the cases (Figure 5.2). In this study, for the upper layer, the combined scheme of B+A+G+I+T produced the highest R^2 (0.65) followed by the scheme of the A+G+I+T dataset (R^2 = 0.63). However, the R^2 values for the middle layer did not increase with the increase in the number of predictor variables when compared the scheme of B+A+G+I+T (R^2 = 0.45) with A+G+I+T (R^2 = 0.46), thus the scheme of A+G+I+T yielded the highest percentage of variations (R^2 = 0.46) for the middle layer. For the lower layer model, the increasing percentage of variation with the B+A+G+I+T scheme (R^2 = 0.44) was not significant when compared with A+G+I+T (R^2 = 0.46). The R^2 values of the combined datasets were mostly greater than the single dataset, for example, R^2 values of spectral indices (I; R^2 = 0.17) and texture features (T; R^2 = 0.25)

were less than those when they were combined with another dataset. Comparing the individual datasets, 48% (R^2 = 0.48) of the variability in the upper layer was described in the model by the spectral indices (I), whereas only 44% (R^2 = 0.44) and 39% (R^2 = 0.39) of the variability was explained by the spectral bands (B) and topographic attributes with geology ancillary data (A+G) respectively. In contrast, the scheme of A+G produced better performance (R^2 = 0.32) than spectral bands (B; R^2 = 0.15) or spectral indices alone (I; R^2 = 0.17) in the lower layer. Thus, the combined topographic attributes and geology with spectral indices dataset (A+G+I) performed better than the spectral indices or A+G dataset to predict the vertical structure of wet eucalypt forests.

Overall, the RF model had the greatest predictive power for the upper layer (R^2 = 0.65) followed by the middle layer (R^2 = 0.46), and the lower layer had the least predictive power (R^2 = 0.44) in wet eucalypt forests. Spectral indices (R^2 = 0.48) and texture features (R^2 = 0.47) were also found useful for predicting the density of the upper layer of the wet eucalypt forests. If the textures or spectral indices are combined with LiDAR-derived topographic attributes, the combined dataset produced greater R^2 and consequently higher model accuracy than either dataset alone.

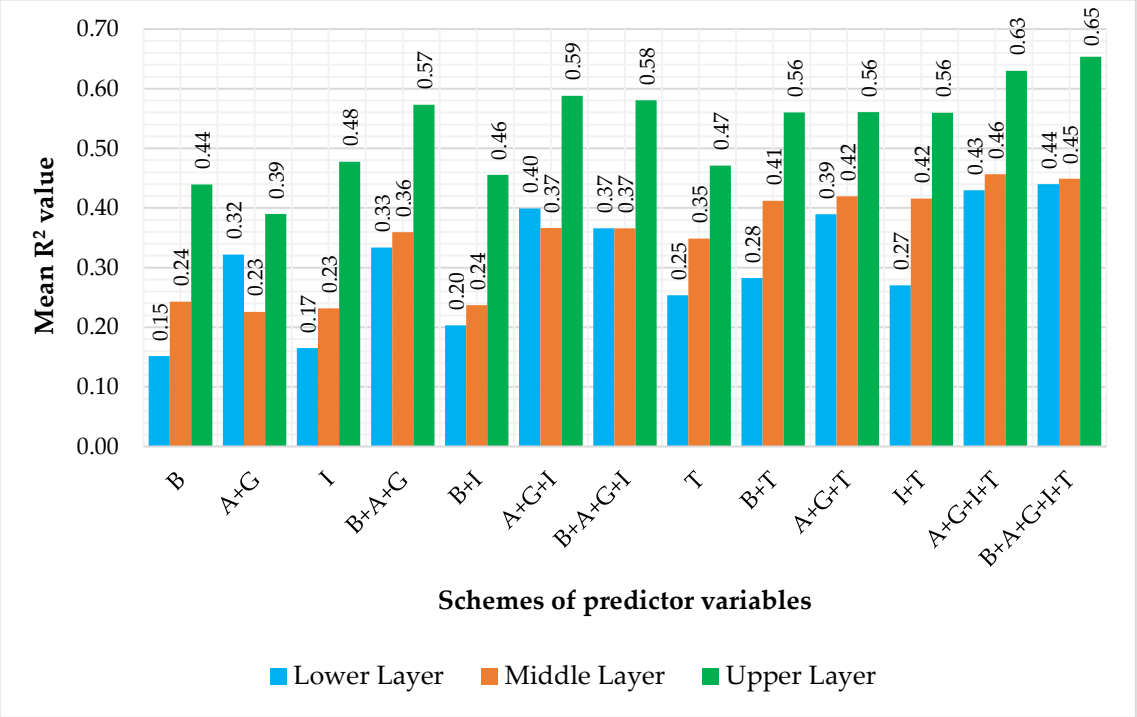


Figure 5.2 Predicted mean R^2 from the 13 schemes of predictor variables derived from the 30 m Landsat-8 (OLI) satellite imagery and the simulated low-density discrete return LiDAR data. Here, a scheme means a group of predictor variables used for RF modelling where spectral bands = B, spectral indices = I, texture features = T, topographic attributes = A, geology vector data = G.

5.3.1.2 Accuracy for simulated 30 m spatial resolution employing WorldView-3 data

The mean R^2 values predicted that 19% to 55% of the variability (i.e. R^2 ranged from 0.19 to 0.55) was explained by the models in the schemes with simulated WorldView-3 dataset (Figure 5.3). As with the Landsat-8 (OLI) data, there was the highest model accuracy for the upper layer, followed by the lower layer, but unlike the Landsat-8 data, the middle layer was the least well predicted in most of the scenarios. Scheme #7 (B+A+G+I) with the combination of the pixel values of 16 spectral bands (B) and their calculated spectral indices (I) with the LiDAR-derived topographic attributes (A) and geology (G) explained the greatest percentage of variance ($R^2 = 0.55$) in the upper layer, followed by that of B+A+G+I+T which also included texture features (scheme #13), and the model using spectral indices (I) had the lowest explanatory value ($R^2 = 0.19$). For the middle layer, the combined spectral indices and texture features (I+T) yielded the highest percentage of variation ($R^2 = 0.38$), whereas the B+A+G+I+T (scheme #13) explained the highest variance ($R^2 = 0.39$) in the lower layer, closely followed by B+A+G ($R^2 = 0.38$; scheme #4).

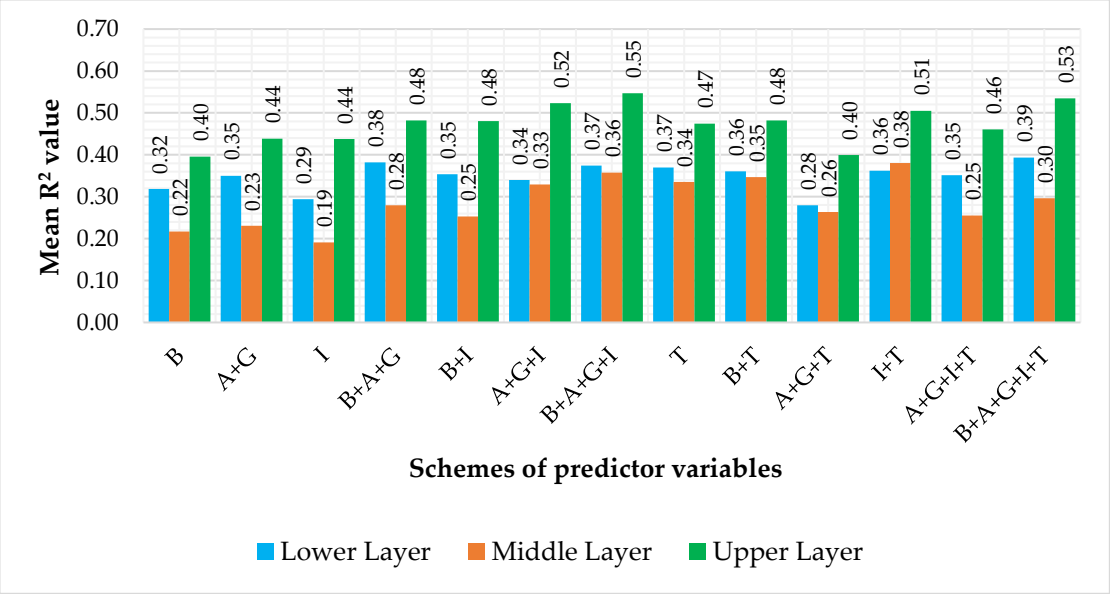


Figure 5.3 Predicted mean R^2 values from the 13 schemes of predictor variables derived from the 30 m simulated WorldView-3 satellite imagery and the simulated low-density discrete return LiDAR data. Here, a scheme means a group of predictor variables used for RF modelling.

Comparing the schemes of the 30 m Landsat-8 (OLI) (Figure 5.2) with the simulated 30 m WorldView-3 datasets (Figure 5.3), the WorldView-3 imagery could predict the lower layer better than the middle layer, whereas the schemes with Landsat-8 (OLI)

dataset could predict the middle layer better than the lower layer. Overall, the Lansat-8 (OLI) dataset performed better (R^2 ranged from 0.15 to 0.65) than the WorldView-3 datasets (R^2 ranged from 0.19 to 0.55).

5.3.1.3 Accuracy for 7.5 m spatial resolution employing the composites of WorldView-3 shortwave infrared and the simulated visible near-infrared data

Considering the thirteen schemes of predictor variables with the simulated WorldView-3 VNIR bands to 7.5 m and SWIR bands of 7.5 m in the RF regression models, the upper layer had the highest predictive power (B+A+G+I+T, $R^2 = 0.32$), followed by the lower (B+A+G+I+T, $R^2 = 0.21$), and the middle layer ($R^2 = 0.14$) had the least predictive capacity. The predicted mean R^2 values ranged from 0.02 to 0.32 (Figure 5.4). When the scheme of a single dataset was compared, the texture features (T) for the upper layer showed the best performance ($R^2 = 0.20$). Combinations of two or more datasets provided better prediction than any one of the individual datasets in this case as well. For the middle layer, the spectral indices alone (I) explained only 3% of the variability in the model ($R^2 = 0.03$), whereas the combined spectral indices and texture features (I+T) explained 10% of the variability ($R^2 = 0.10$).

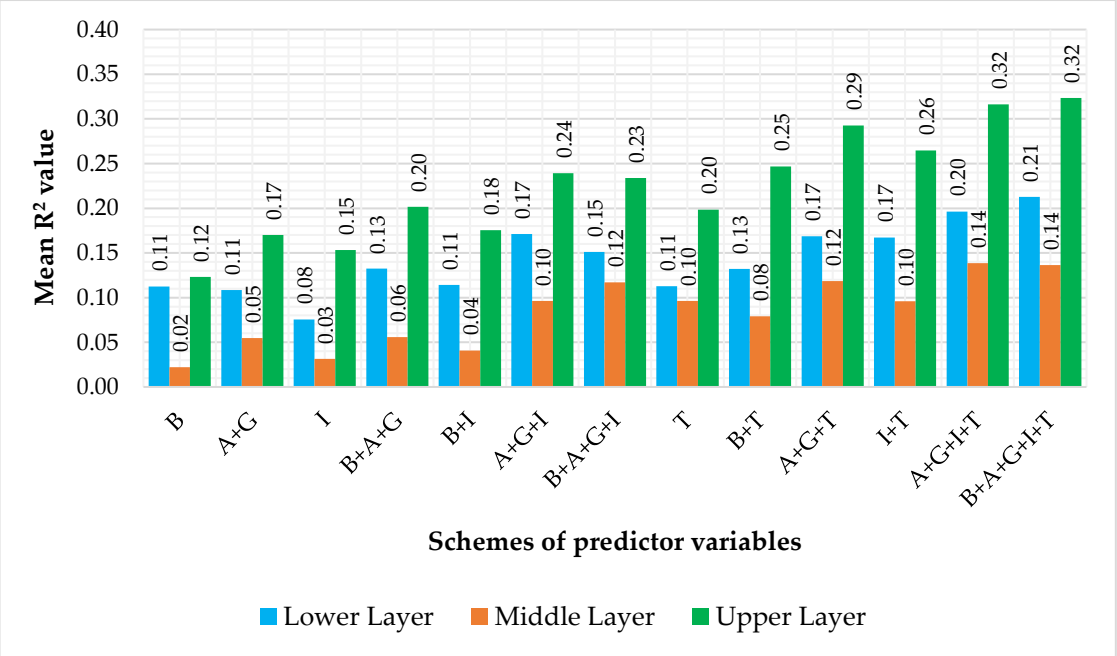


Figure 5.4 Predicted mean R^2 from the 13 schemes of predictor variables derived from the 7.5 m simulated VNIR and original SWIR bands of WorldView-3 satellite imagery and the simulated low-density discrete return LiDAR data. Here, a scheme means a group of predictor variables used for RF modelling.

5.3.1.4 Accuracy for 1.6 m pixel spatial resolution employing WorldView-3 data

Considering the thirteen schemes of derived variables from the original VNIR eight bands of WorldView-3 imagery (pixel size of 1.60 m) and LiDAR-derived topographic attributes, the overall result demonstrated poor predictive power with mean R^2 values ranging from 0.0 to 0.08 (Figure 5.5). There was a stronger predictive capacity for the upper layer compared to the middle and lower layers. B+A+G+I+T, A+G+I+T, and I+T had a similar result and the highest R^2 value of 0.08 for the upper layer. The scheme of B+T provided the highest R^2 value of 0.04 for the lower layer, and the scheme of A+G+I+T had the highest R^2 value of 0.02 for the middle layer if compared within those thirteen schemes of datasets. Thus, it could be inferred that the WorldView-3 imagery with the pixel size of 1.60 m and their combinations of the predictor variables were not appropriate for predicting the vertical structure of wet eucalypt forests.

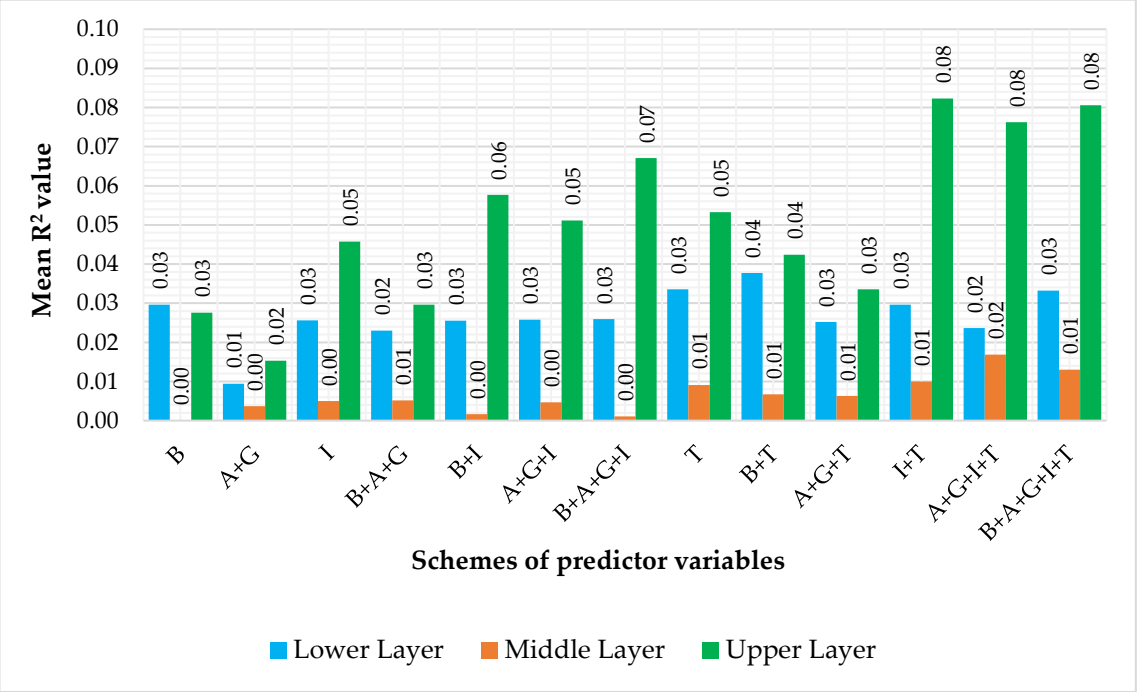


Figure 5.5 Predicted mean R^2 from the 13 schemes of predictor variables derived from the 1.60 m VNIR bands of WorldView-3 satellite imagery and the simulated low-density discrete return LiDAR data. Here, a scheme means a group of predictor variables used for RF modelling.

5.3.2 Model validation

This study evaluated the predicted model accuracies using an independent validation dataset to confirm the robustness and transferability of the acquired models (see Figure 5.6 to Figure 5.9). In this study, the differences between the predicted and validated R^2

values (validated–predicted) were small (<5.67%) across all models. In the figures, positive values indicate more and negative values less than the validated accuracies. This study considered two criteria for transferability: (1) the model with the higher prediction and validation accuracies, and (2) the smaller the difference between validated and predicted R^2 values, the more the transferability. This study indicates the developed models can be transferred to other areas in this forest landscape.

When the percentage variations of predicted models compared with the validated models, the differences fluctuated according to the deployed schemes of the Landsat-8 (OLI) data with topographic attributes and geology data. The differences in percentage variation between the predicted and validated models ranged from -2.53% to +3.32% and from -3.44% to +1.72% respectively in the lower and middle layers (Figure 5.6). The validated model explained the highest variation of 63% ($R^2 = 0.63$) for the upper layer in the scheme of Landsat-8 (OLI) with a topographic dataset with a difference of -0.26% if compared with the percentage variation of the predicted model. Likewise, the validated model of B+A+G+I+T explained 62% of the variation in the upper layer, which was 3% lower than the predicted variation of the same scheme. Another scheme with the combined pixel values of spectral bands and spectral indices (B+I) yielded a validated model with 49% of the variations ($R^2 = 0.49$) which was 3% greater in comparison with the predicted variation of 46%. The other schemes of datasets produced the variability differences of up to ±2%. Overall, in this study, the A+G (+0.01%), followed by the A+G+I+T scheme demonstrated the most robust and transferability capability among the thirteen schemes for the upper layer datasets (-0.26%). The scheme for texture features (T) proved the most robust for both the middle layer (+0.03%) and the lower layer (-0.03%) of wet eucalypt forests.

When assessing the 30 m simulated WorldView-3 dataset, the variation differences between predicted and validated models ranged from -4.29% to +3.44% for the upper layer, from -5.67% to +4.58% for the middle layer, and from -5.67% to +1.61% for the lower layer (Figure 5.7). In this study, the B+A+G+I+T scheme produced the least percentage of difference, i.e. +0.006% in the lower layer, whereas that with B+A+G+I scheme had +1.25%. Thus, it can be inferred that the B+A+G+I+T (scheme #13) had the strongest prediction capability for the lower layer followed by the A+G+I+T (scheme #12) (-0.21%). Considering the differences between prediction and validation accuracies, the scheme of texture features and spectral indices (I+T) with the simulated 30 m WorldView-3 dataset showed better performance than that with LiDAR-derived topographic attributes for the upper layer that indicated the derivatives of WorldView-

3 dataset could be deployed to predict the upper layer of wet eucalypt forest, excluding LiDAR-derived topographic attributes. For the B+A+G+I+T with the 7.5 m WorlView-3 datasets, the differences in the percentage variations between the predicted and validated models ranged from -2.16% to +3.61% (Figure 5.8). Similarly, the variation differences in the validated model with the schemes of 1.60 m WordView-3 dataset ranged from -1.53% to +1.72% for the upper layer, from -0.65% to 0.93% for the middle layer and from -1.54% to +1.42% for the lower layer (Figure 5.9).

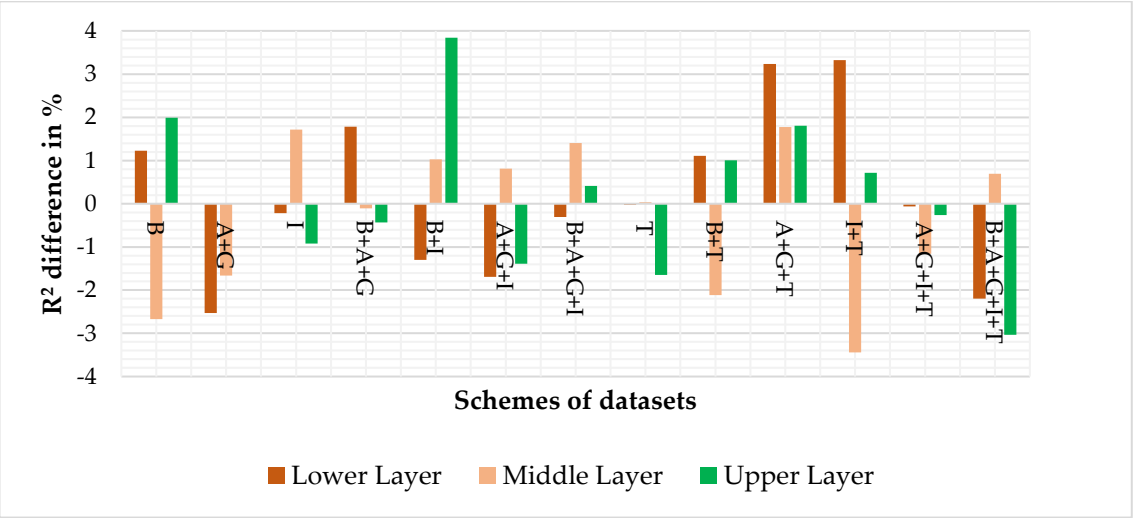


Figure 5.6 Accuracy differences (validated-predicted) of thirteen schemes using the 30 m Landsat-8 satellite imagery with topographic attributes and geology data. Here, Spectral bands = B, spectral indices = I, texture features = T, topographic attributes = A, geology vector data = G.

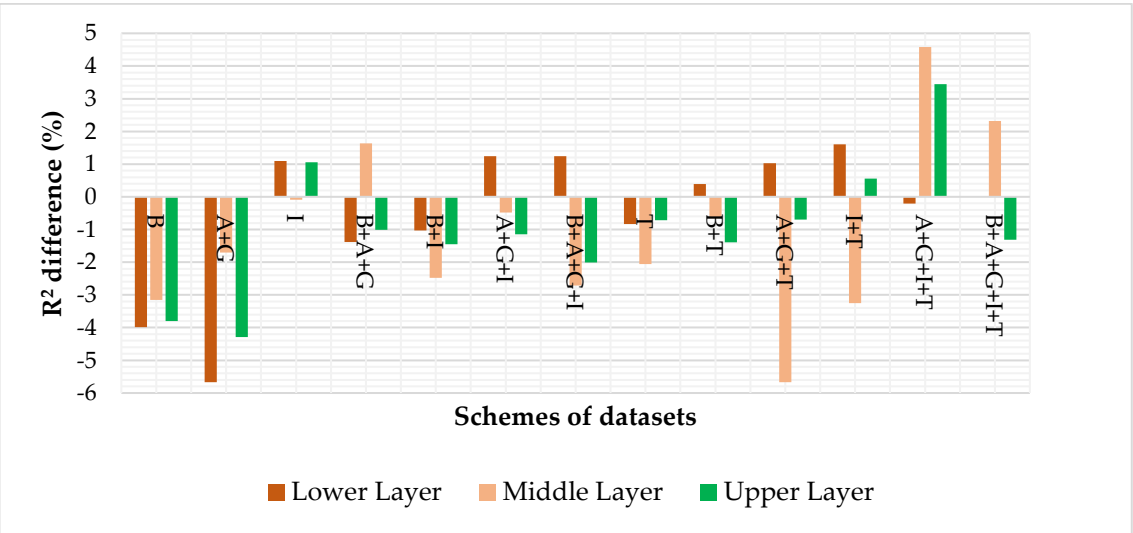


Figure 5.7 Accuracy differences (validated-predicted) of thirteen schemes using the 30 m WorldView-3 imagery with topographic attributes and geology data.

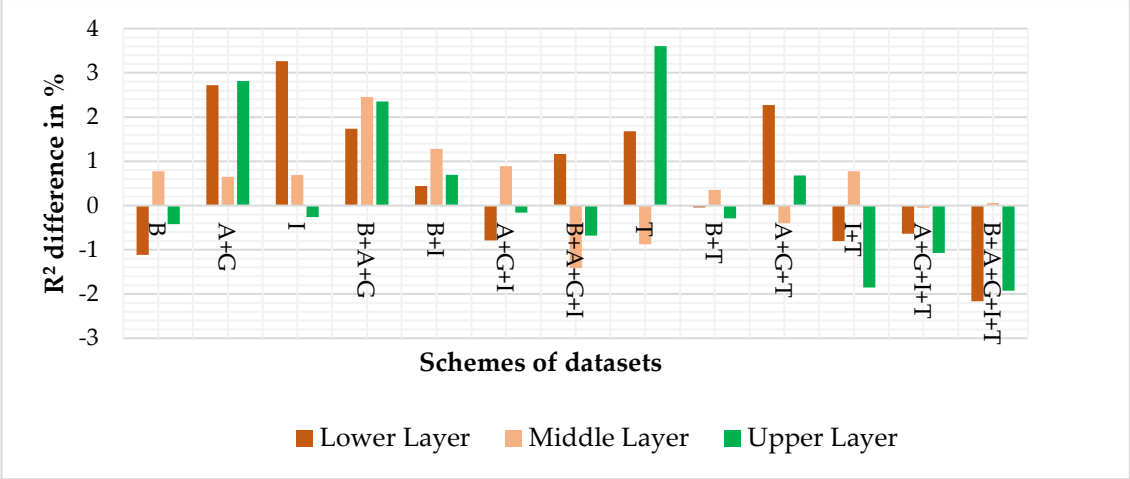


Figure 5.8 Accuracy differences (validated-predicted) of thirteen schemes using the 7.5 m simulated VNIR and original SWIR bands of WorldView-3 imagery with topographic attributes and geology data.

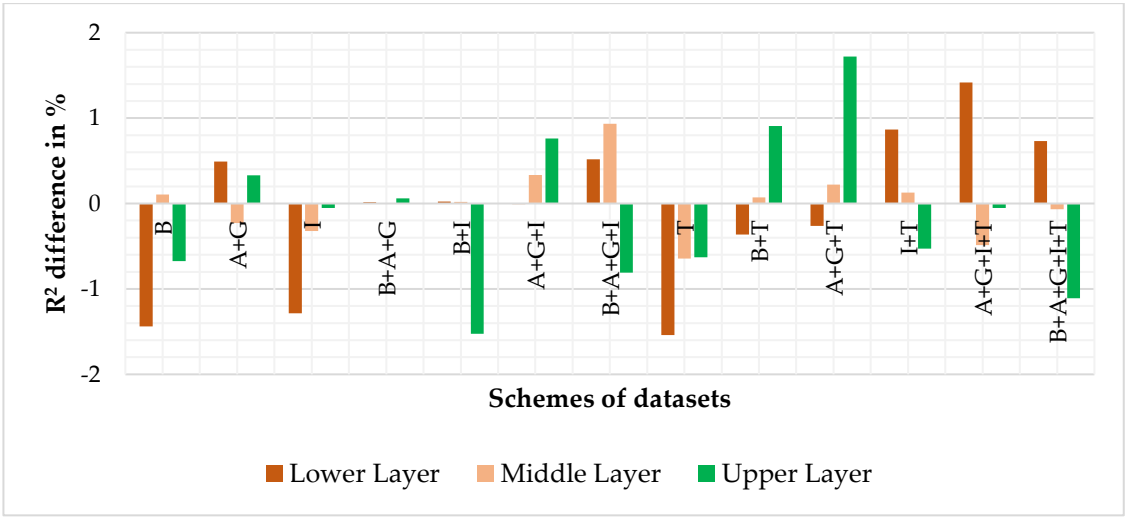


Figure 5. 9 Accuracy differences (validated-predicted) of thirteen schemes using the 1.60 m VNIR bands of WorldView-3 satellite imagery with topographic attributes and geology data.

Considering the pixel sizes and the predictor variables used, the 30 m pixel size showed better performance than the pixel sizes of 7.5 m and 1.60 m. When the simulated 30 m WorldView-3 data was compared with the 30 m Landsat-8 (OLI) dataset, the Landsat-8 (OLI) dataset produced higher accuracies in the cases of prediction as well as validation models. The 30 m WorldView-3 imagery was resampled from the original 1.6 m VNIR 8 bands and 7.5 m SWIR 8 bands of the WorldView-3 imagery and that might affect the properties of the imagery, whereas the 30 m Landsat-8 OLI surface reflectance imagery was provided at the sensor-specific resolution. The resampling of WorldView-3 data has most likely led to a decrease in data quality, and hence a decrease in the model accuracy. However, this could be an

interesting topic for further research. LiDAR-derived topographic attributes contributed significantly in the case of 30 m datasets, but that was negligible or could not contribute to the models developed using the pixel sizes of 7.5 m and 1.60 m datasets. The selected spectral indices and texture features from Landsat-8 (OLI) imagery with the simulated low-density discrete return LiDAR-derived topographic attributes are most suitable to predict the vertical structure of the wet eucalypt forests.

5.4 Discussion

With a view to presenting cost-effective remote sensing approaches for modelling vertical forest structure of three canopy layers in a wet eucalypt forest, this chapter deployed multiple variables derived from freely available medium resolution Landsat-8 (OLI) surface reflectance and commercial high-resolution WorldView-3 satellite imagery. These datasets were combined with the topographic attributes extracted from simulated low-density discrete return LiDAR data. A key novelty of this research lies in the fusion of the derivatives of the three different data sources and assessment of the optimum spatial resolution suitable for modelling vertical forest structure (i.e. canopy density layers) using a random forest regression technique. Previous studies used fewer variables and either spectral indices or texture features from a single source of multispectral satellite data alone or with LiDAR data. This research assessed the strengths of the simulated operational LiDAR data relative to the high resolution (1.60 m) and medium resolution (30 m) datasets. This study also discussed how aggregation schemes of the high-resolution WorldView-3 dataset perform in the modelling of vertical forest structure.

5.4.1 Comparison of model accuracies in the wet eucalypt forest

Comparing all the thirteen schemes, the spatial resolution of the freely available 30 m Landsat-8 (OLI) dataset and its combinations with the topographic attributes derived from the simulated operational LiDAR and a geology data produced the best result (R^2 = 0.65) for the upper layer followed by the middle layer (R^2 = 0.46) and then the lower layer (R^2 = 0.44) (Figure 5.2). There was generally an increase in the predictive accuracy with the increasing numbers of predictor variables, and so the model that included pixel-values of spectral bands (B), topographic attributes (A), spectral indices (I), texture features (T) and ancillary geology vector data (G) produced the best overall prediction accuracy (R^2 = 0.65) for the upper layer. The next best results after those for Landsat-8 with the topographic attributes and a geology data were produced by the 30 m simulated WorlView-3 dataset with the topographic attributes and geology (R^2 =

0.53) for the upper layer (Figure 5.3). The addition of topographic attributes derived from the simulated operational LiDAR data and a geology layer combined with the multispectral datasets provided higher model performance than the combinations with the spectral indices and texture features in isolation. In this study, the spatial resolution of the 1.60 m WorldView-3 dataset was not considered appropriate for modelling vertical forest structure as the R^2 values ranged from 0.0 to 0.08 (Figure 5.5). The results demonstrate that the resolution of the remote sensing dataset is directly proportional to the prediction accuracy of understory canopy layers, which is why the 1.6 m WorldView-3 dataset performed very poorly with an R^2 value of 0 to 0.08 and the 30 m WorldView-3 dataset performed the higher model accuracy (R^2 value of 0.55). Very small structural elements in the forest (smaller than 1.6 m) are simply not representative for the scale of the classes under consideration. Regarding the rationale for expecting a 2D data to estimate a 3D variable, the (extensive) remote sensing literature in radiative transfer modelling indicates that the spectral signal in satellite image pixels is not only influenced by 2D spatial variability, but also by the 3D structure of vegetation. The way that photons are absorbed, transmitted, reflected and re-scattered by leaves and branches results in a complex signal in 2D image pixels.

The above results were in broad accordance with those of Latifi et al. (2012b) who used Landsat-5 thematic mapper (TM) imagery with LiDAR data and predicted forest attributes using twelve predictor variables. They also used RF models and reported that accuracy was increased with the number of predictor variables increased. They also emphasized the dominance of topographic variables derived from LiDAR data over the multispectral satellite data for predicting forest structural attributes. Considering spatial extents limited to the local areas, Wing et al. (2012) reviewed the different studies related to understory layer models and reported that R^2 values ranged from 0.20 to 0.45. Compared with their outputs, this study produced better results. The 1.60 m WorldView-3 data produced the least accuracy. Cohen and Spies (1992) postulated that the pixel size of Landsat TM data is roughly equivalent in size to the tree crown, and even complex stands contain crowns larger than one pixel, whereas a large tree may appear in many pixels if pixel sizes are smaller, for example, 7.5 m or 1.60 m, and this consequently produced lower model accuracy. Azaele et al. (2012) reinforced this view and reported that they could better measure an object that occupied a pixel. Woodcock and Strahler (1987) concluded that the spatial structure of the image is the function of spatial resolution. The interests of the investigators, therefore, determine the optimal pixel sizes: a study requiring information on individual canopy trees may be expected to require a larger pixel size than studies

requiring information on individual branches or smaller plants in lower canopy layers. Similarly, Zald et al. (2014) exemplified that when the LiDAR-derived forest metrics were extracted for the canopy of tree plots, the overhanging branches from trees outside the plots influenced the forest metrics. Thus, the ability of LiDAR metrics to represent the trees within plots was reduced as the plot size decreases. Kamal et al. (2014) applied original and resampled WorldView-2 and demonstrated that a pixel size ≤ 2 m was appropriate for mapping inter-canopy features, such as canopy gaps, and a pixel size ≥ 4 m was more suitable for mapping vegetation formation and communities in mangrove forests. Their prediction accuracy improved as the spatial resolution increased. Dash et al. (2016) also recommended RapidEye data compared to WorldView-2 and IKONOS; however, they believed that such outcomes might depend on different factors, such as forest type and age, size and quality of datasets. They did not find any tangible improvement in a model performance combining LiDAR data with RapidEye data and compared with the models using LiDAR metrics alone. They commented that the outputs depended on the objective of the modelling whether they were forest type, land cover or species classification where RapidEye spectral information might be more useful. They concluded that LiDAR was the most valuable data source for predicting forest structure attributes, while RapidEye data could improve the prediction accuracy by a negligible amount. This contrasts with my results where Landsat-8 (OLI) data combined with operational resolution LiDAR derivatives produced the best accuracy for describing the density of the three canopy layers. It is important to note that I could not compare direct canopy density estimates for operational LiDAR with those of high resolution LiDAR. Operational LiDAR was unavailable for much of the study area, and it was deemed inappropriate to compare direct density estimates from the down-sampled high resolution LiDAR since this subset would have included exactly the same points as for the comparison dataset, artificially inflating their explanatory value. This study therefore focused on the combination of multispectral satellite data and topographic data derived from the simulated operational LiDAR only.

Considering the published literature and my results, it was inferred that the scaling up from small areas to larger areas are relevant for predicting the vertical forest structure over large landscapes. It is noteworthy here that surface reflectance data are the key to the mapping of forest stand structure containing the reflectance of overstory canopies. Nowadays, Landsat surface reflectance data is freely available to the public as an analysis-ready data product by the USGS that opened opportunities for mapping large geographical areas with high-quality optical products (Matasci et al. 2018). My results

demonstrate the potential value of these datasets, with improved overall results compared to more expensive Worldview-3 data.

In the scheme of Landsat-8 (OLI) combinations, spectral indices and texture features respectively yielded R^2 values of 0.48 and 0.47 for the upper layer modelling. For the middle and lower layers, the Landsat-8 (OLI) dataset produced the best overall results for each canopy layer. However, the 30 m Landsat -8 (OLI) and the 30 m simulated WorldView-3 datasets produced contrasting outcomes for relative abilities to predict the middle and lower canopy layers. The original and simulated WorldView-3 data could predict lower layers better than the middle layers, whereas Landsat-8 (OLI) data predicted middle layers better than lower layers. The studies of Azaele et al. (2012) and Cohen and Spies (1992) found that the spatial resolution of the dataset should match the object to be predicted. Therefore, the model outputs obtained by the scheme of Landsat-8 (OLI) combinations were satisfactory. While comparing the R^2 values of the combined Landsat-8 (OLI) data with topographic attributes and a geology data for the middle layer (R^2 values ranged from 0.23 to 0.46; Figure 5.2) with those of the 30 m simulated WorldView-3 dataset with topographic attributes and a geology layer (R^2 ranged from 0.19 to 0.38; Figure 5.3), the scheme of the Landsat-8 (OLI) datasets with topographic attributes again had better predictive power. Overall, the LiDAR-derived topographic attributes and geology data combined with the Lansat-8 (OLI) dataset produced better model performance than those with the scheme of WorldView-3 datasets. Also, WorldView-3 images are expensive. This study recommends that when high-density LiDAR data are unavailable, using freely available Landsat-8 imagery combined with DTM derivatives (topographic attributes), texture measures and vegetation indices can be adequate for forest managers and planners to assess the vertical forest structure of wet eucalypt forests, to contribute to the sustainable management, planning, and monitoring of forests.

In general, the predicted accuracies of the models compared favourably with those of previous research in other forest systems and confirmed the validity of the approach of this study. A study conducted by Zald et al. (2016) advocated that prediction accuracy depended on structural variable type using derivatives from LiDAR and Landsat data, and presented R^2 ranging from 0.0 to 0.77 with the highest value for structural variables including live trees, intermediate values for live tree density, lower values for snag density and the lowest for downed wood. They found that metrics derived from Landsat data improved the prediction accuracies of all the structural variables, although those improvements were less than the inclusion of LiDAR data. Maselli et al.

(2005) suggested that spectral differences depended on variations in the forest composition, canopy density, topography and understory conditions. Similarly, Wallner et al. (2014) stratified forest plots based on forest types and showed R^2 values ranging from 0.37 to 0.63 for modelling stand structural attributes using RapidEye data (5 m spatial resolution), particularly R^2 values of 0.4 for stand density. Kayitakire et al. (2006) reported the model prediction with an R^2 value of 0.38 for stand density. Therefore, this research can act as a benchmark for including a broad range of explanatory variables, however, further research is required to translate this approach to other types of forest ecosystems with different topography and understory conditions.

5.4.2 Comparison of model validations in the wet eucalypt forest

Validating the predicted accuracies provided promising goodness of fit statistics that could be applicable to the larger geographical areas (Matasci et al. 2018). The difference between model accuracies and validation accuracies was very low for WorldView-3 dataset combinations ranging from +1.72% to -1.54% and most of the schemes had less than ± 1.0% difference, but since the overall explanatory power of the 1.6 m and 7.5 m WorldView-3 data was so low, this study does not recommend these data for predicting forest structural information.

Comparing the model validation results of this research with other studies, Dash et al. (2016) evaluated their study with the models and found the reduction of relative root mean square error (RRMSE) of 1.6% over all dependent variables (predicted vs validated accuracy). Ahmed et al. (2015) reported that disturbance information could be applied to improve the estimation of stand level canopy structure and validated their study, for example, they achieved the best validation results with mature forest (R^2 = 0.88 and bias = 0.16 m for canopy height) using random forest modelling. Although satellite data were, in general, valuable to characterise forest structural variables in a variety of environments (Gebreslasie et al. 2010), their limitations, for example, Landsat vegetation indices in the dense forests contributed to the underestimated prediction of mature forest attributes and productive stands due to saturation in the closed canopy with high biomass (e.g., Avitabile et al. 2012; Avitabile et al. 2011; Bolton et al. 2018; Gasparri et al. 2010; Zald et al. 2016). Bolton et al. (2018) achieved negative model bias (-1.2 to -2.1%) and validated their research to new blocks to examine transferability. They suggested that the structural variability in the new blocks might not have been captured in the training plots used in the model development, and the bias varied from -3.9% to –8.0%. In this study, the RF model was

able to predict vertical forest structure with the sampled training and test datasets using LiDAR, and generate maps in a broader scale using remote sensing optical imagery that demonstrated the transferability of the models to estimate at the scale of interest for forest managers. However, further research would be required to test the transferability to other wet forest landscapes.

From the aforementioned discussion, the 30 m Landsat-8 (OLI) imagery and its combined schemes were able to predict the vertical forest structure (i.e. canopy density) of different layers (lower, middle and upper) better than the other schemes of WorldView-3 datasets. The topographic attributes derived from a DTM could be deployed in combination with the Landsat-8 (OLI) dataset, as their contributions were high in the models. The freely available Landsat surface reflectance dataset could be used for modelling and mapping the larger geographical areas that could be easily adopted by researchers and forest managers.

5.4.3 Implications of the study

This study is important for forest managers and researchers who mostly use field data for training and validation purposes. This study applied high-resolution LiDAR data to represent forest structural composition and showed that satellite remote sensing data could be applied for training and validation purposes and could predict and map forests at various scales without requiring field data. A novel aspect of this study was using simulated LiDAR, Landsat-8 and WorldView-3 datasets to compare different remote sensing products and pixel sizes. The contribution of the simulated operational LiDAR-derived topographic attributes was high, so this study exemplified that low-density discrete return LiDAR data could be used to predict and map the vertical forest structure. Future research should compare these results to those for operational resolution LiDAR, although the expense of data acquisition would need to be considered. This data can, therefore, contribute to sustainable management, planning, and monitoring of the wet eucalypt forests (van Galen et al. 2018). Although this research was limited to predicting vertical forest structure across three scales, understanding the effects of spatial scales of vertical forest structure will be helpful for the multi-scale management of wildlife habitats (animals, birds, and insects) to sustain forest-dependent species.

LiDAR-based models of the vertical forest structure could be useful for resource managers to update the spatial forest inventories based on more recent datasets (Fekety et al. 2014). This approach could be applied across remote, inaccessible and large areas to map forest attributes using remotely sensed datasets on vegetation

structure (Zald et al. 2016). Detailed spatially explicit maps of forest structure are increasingly being required to support science, policy and reporting purposes where synergistic use of optical satellite imagery and LiDAR data might be useful for mapping the vertical forest structure across the wet eucalypt forests (Matasci et al. 2018). Modelling multivariate structural complexity of a forest would allow forest managers to map local structural complexity variations across the entire forest (Pasher and King 2010). This can also give information on damage and disturbances to forest-dwelling organisms and their diversity, which has become of great interest to forest practitioners involved in forest inventories (Winter et al. 2008). The importance of understories and their contribution to vertical vegetation layers is underscored by positive relationships of habitat diversity/heterogeneity with biodiversity elements. The effects of habitat heterogeneity may differ with respect to the spatial scale, thus understanding the vertical forest structural composition has profound implications for the nature conservation and biodiversity management (Tews et al. 2004).

5.5 Conclusions

A key advance of this research was the integration of multivariate derivatives from three different data sources to present an optimum spatial resolution of the datasets appropriate for modelling vertical forest structure (i.e. canopy density) deploying random forest machine learning. This study validated the developed models using independent datasets to confirm their robustness and transferability. The following conclusions can be made:

- The fusion of the derivatives from the 30 m Landsat-8 (OLI) satellite imagery and DTM derivatives (topographic attributes) produced the best overall results (ranging R^2 values from 0.15 to 0.65). Therefore, the schemes of the Landsat-8 (OLI) datasets were robust to predict the vertical forest structure of three canopy layers (lower, middle and upper) in comparison with three high-resolution WorldView-3 dataset schemes. Landsat-8 data is freely available high-quality data so forest managers and planners could easily adopt the outputs, even where LiDAR data are unavailable.
- The differences between the predicted and validated accuracies were less than 5.67% for all models, indicating the developed models could be transferred to similar forests outside of the focal research landscape.
- Spectral indices and texture features were useful for predicting the upper layer of the wet eucalypt forest. If the texture features or spectral indices combined with

topographic attributes, the combined dataset produced higher model accuracy than either dataset alone.

- The analysed output demonstrated that the simulated 30 m pixel size of WorldView-3 data had better performance than the pixel sizes of 7.5 m and 1.60 m. The schemes with WorldView-3 datasets with the pixel size of 1.60 m were not appropriate for predicting the vertical structure of the wet eucalypt forest.
- The results showed an increase in the model accuracies with an increasing number of predictor variables in most of the schemes, illustrating the merit of combining topographic attributes, geology, texture features, and spectral indices.

Further work would be required to extending the results to different forest types and age classes since this study was confined to mature wet eucalypt forests. It would be very valuable to test this approach integrating the topographic attributes derived from the recently launched space-borne Global Ecosystem Dynamics Investigation (GEDI) LiDAR data (footprints averaging 25 m in diameter) with the 30 m Landsat-8 (OLI) imagery for predicting and mapping large geographic areas.

Finally, this approach could be applied where there are limited resources for field-based research as remote sensing data could be used for training and validation purposes from small scales to broad scales. Thus, this research addresses data complexities, including multidimensionality and nonlinearity in multisource data, and provides a robust approach for the assessment of wet eucalypt forest structure.

Chapter 6

Synthesis and Conclusions

6.1 Synthesis

This research was conducted with the aim to develop approaches for assessing and mapping tree species distribution and forest structure using multi-source remote sensing data for Tasmanian wet eucalypt forests that contained tall eucalypts over dense understories of rainforest and wet sclerophyll species. The first case study segmented tree crowns and classified both the overstory and understory species based on objects using highly complex raw hyperspectral imagery and LiDAR point clouds. This research should be helpful for planning sustainable management of wet eucalypt forest biodiversity and monitoring changes in species distribution. The composition of tree and shrub species is related to the vertical structure of forests and so the second study exemplified relationships between the topographic attributes with understory forest structure. This study solely depends on LiDAR remote sensing and geological auxiliary data to assess the over- and understory densities of wet eucalypt forests. The results provided an in-depth insight into the estimation of three layers of wet eucalypt forests without the use of field inventory data and could be tested for other types of forests. The understory forest characteristics play a key role in explaining the variations of fuel loads and biomass/carbon (Temesgen et al. 2015). The third case study was designed to broaden the scope of topographic attributes derived from LiDAR data (Chapter 4). This study deployed an extensive number of variables (texture features and indices) from the WorldView-3 and Landsat-8 OLI remote sensing data in addition to the operational LiDAR data to assess robustness for predicting vertical structure across multiple scales. This study found that the fusion of the derivatives from the 30 m Landsat-8 OLI data with topographic attributes was robust and useful for predicting the vertical structure of wet eucalypt forests. Although Landsat-8 OLI data provided the best result, the two-dimensional data combined with three-dimensional LiDAR data provided an improved solution. The operational LiDAR data still proved useful and is substantially cheaper to obtain compared to the high-density LiDAR data over large geographical areas. However, these approaches and datasets could be further developed to assess the aboveground carbon and the basal area of commercial timber. Thus, the domain of the studies could be extended to beyond just predicting canopy cover or over-and understory and designed more independently to address the caveats stated within this PhD research.

6.2 Conclusions

The overall aim of this thesis was to develop remote sensing approaches for the assessment and mapping of tree species distribution and vertical forest structure using remote sensing datasets of different sources in wet eucalypt forests of Tasmania, Australia. The motivation of this research work began with recognition of the limitations of the traditional technique of aerial photography mapping (PI-type maps) for species or community classification and stand structure in Tasmanian forests. Because canopy eucalypts obscure lower vegetation layers, it is difficult to use aerial photography to predict understory vegetation composition and structure. However, some remote sensing techniques can overcome these difficulties. Thus, forest managers are looking to remote sensing approaches for cost-effective solutions that enable coverage of remote regions and large geographical areas.

The research described in this thesis tested different scenarios of data fusion on different pixel sizes using a large number of variables to propose suitable datasets, and optimal scales and pixel sizes for describing forest composition and structure. Thus, this research work focused on assessing the capability of airborne LiDAR data and hyperspectral imagery for object-based tree crown segmentation and species classification, and for predicting vertical forest structural composition. This research aimed to identify key factors that contribute to the accurate mapping of forest species distribution and vertical structure.

This final chapter presents a summary of the findings of the whole study and highlights the approaches and their applications. This chapter identifies the limitations and future directions of the research and concludes with final remarks.

6.2.1 Object-based tree crown segmentation and species classification

Objective 1 (Chapter 3)

To assess the capability of airborne LiDAR and hyperspectral data for optimal tree crown segmentation and species classification in wet eucalypt forest. Additionally, this research explored whether understory species could be identified with a multi-source dataset and presented the contribution of a LiDAR-derived Canopy Height Model (CHM) to improve accuracy in the object-based classification.

Forest managers and planners around the world have been using multispectral satellite data for more than four decades for forest species classification and identification purposes. However, the approach has limitations; namely saturation where there are

high biomass and information about the understory vegetation which is typically obscured by canopy vegetation. Several studies have deployed airborne hyperspectral imagery and LiDAR data separately, but only a few studies have fused these two datasets for tree crown segmentation and species classification. Chapter 3 in this thesis, therefore, utilised different combinations of airborne LiDAR-derived CHM and hyperspectral imagery and explored an approach of optimal segmentation parameters to obtain the highest segmentation accuracy compared to digitized tree crowns within a 1.6 ha permanent research plot. Because exploring optimal segmentation parameters is still a research challenge in order to accurately classify forest tree species based on objects, this research deployed a multiresolution segmentation technique and tested various options for the scale parameter (5 to 45) and for the weightings of shape (0.05-1.0) and compactness parameters (0.05 to 1.0). The objected-based tree species classification was performed using a random forest machine learning approach.

The fused Minimum Noise Fraction (MNF) and CHM dataset in this study provided the highest segmentation accuracy (88.71%) with scale parameter, shape and compactness of 10, 0.3 and 0.2 respectively. Overall the fused dataset of hyperspectral imagery, CHM and indices provided the highest classification accuracy (OA = 66.67%, k = 0.52) followed by the fused dataset of MNF and CHM (OA = 66.04%, k = 0.49). HSI dataset alone yielded the lowest classification accuracy (OA = 59.02%, k = 0.30). While considering particular species, the classification accuracy for particular species varied from species to species and also depended on how many species classes were included. For example, accuracy for *Eucalyptus obliqua* was 90.86% when four vegetation classes were included and dropped slightly to 86.11% for five classes. An important understory species, *Dicksonia antarctica* was classified with an accuracy of 83.54% and 84.64% if we classified the forest stand into four and five classes respectively.

Thus, the research in Chapter 3 explored the optimal segmentation of tree crowns and classified both the overstory and understory species with sufficient accuracy to suggest that the approach could be useful for comprehensive forest planning and monitoring purposes. The fused LiDAR and hyperspectral datasets were more robust than hyperspectral or LiDAR data alone for spatially discriminating and classifying wet eucalypt forests to species level using object-based classification, illustrating the merits of integrating LiDAR data with hyperspectral imagery. Whenever the number of tree species was increased, the classification accuracy was decreased, and we were only able to do this for the most common species in the plot, meaning that information on potentially important but less dominant species would be lacking with this approach.

This suggests the approach may be best suited to mapping forest stands with relatively a few dominant plant species, thus the classification of 4 to 5 classes is relatively limited even for such high-quality datasets with high-resolution hyperspectral and LiDAR data. The maps of this study serve as an important step towards the forest planning and monitoring to facilitate the sustainable management of wet eucalypt forests and the conservation of their plant biodiversity.

6.2.2 Prediction of the density of understory layers

Objective 2 (Chapter 4)

To predict the density of three understory layers of a wet eucalypt forest using topographic attributes and geological data. Additionally, this study explored optimal parameters for the random forest regression algorithm and identified the most appropriate spatial resolution for predicting understory layers. This study also assessed the relative importance of various input variables.

Predicting understory structural properties and understanding the interaction between forest structure and topographic attributes and geology is relevant to the management and conservation of a forested wet eucalypt ecosystem for high socio-economic value. These forests in Tasmania are labour intensive to map using aerial photography and other traditional techniques, which have limited capacity to describe sub-canopy vegetation layers. Characterizing understory structural properties in complex forests with dense canopies is particularly challenging, and thus the topic of this PhD research. This research utilised LiDAR-derived topographic attributes and a geology vector map to develop a model and map three understory layers using five different pixel sizes in a mature wet eucalypt forest landscape.

In Chapter 4 of this thesis, high-density airborne discrete return LiDAR data were processed and deployed to derive density estimates for three canopy layers and a LiDAR-derived DTM was used to derive predictor variables with five different pixel sizes. This research used random forest regression for the prediction of understory layers. Based on variable importance, five variables were selected to illustrate their effects on the prediction of three canopy density layers and finally generated maps of the predicted understory layers. The results of this work showed that the 30 m resolution models performed best overall as the upper layer, middle layer and lower layer provided MRMSEs of 7.61%, 9.76%, and 8.76% respectively, whereas those with MRMSEs of 17.82%, 20.81% and 25.92% for the 1 m pixel size respectively. Those results indicated the 10.21%, 11.05% and 17.16% information for the upper, middle and

lower layers respectively was lost from the pixel size of 30 m to 1 m. The upper layer had the highest prediction accuracies with the mean value of the root mean square errors (MRMSE) of 7.61% (R^2 = 0.82), and the middle layer produced the lowest predictive power (MRMSE = 9.76%, R^2 = 0.77) with the training dataset. While comparing the variable importance scores, geology provided the highest variable importance score for four resolutions (i.e. 5 m, 10 m, 20 m and 30 m) and terrain position index produced the highest variable importance with %IncMSE value of 37.36% for 1 m resolution.

This research demonstrates that LiDAR-derived topographic attributes and a geology vector map are useful predictors indicating significant relationships with the understory vegetation structure in wet eucalypt forests. The novelty of this research is that understory layers can be predicted from a DTM even if field data or high-density LiDAR data are not available. However, this chapter demonstrated that topographic derivatives were not sufficient to get a good model prediction of understory structural layers of a wet eucalypt forest.

6.2.3 Assessing multispectral satellite imagery for predicting vertical forest structure

Objective 3 (Chapter 5)

To assess the robustness of multispectral satellite imagery with or without LiDAR topographic attributes to predict vertical structure across multiple scales in wet eucalypt forests using random forest regression modelling. Additionally, this research explored the optimum resolution of the datasets and examined the capability for the fusion of derivatives (vegetation indices, texture features, and topographic attributes) for improving the prediction accuracy of vertical forest structure. This research work utilised derivatives from three remote sensing datasets, i.e. Airborne LiDAR, WorldView-3 and Landsat-8 (OLI) datasets.

Measuring vertical forest structure in the field is a challenge for forest managers, and almost impossible in areas that are inaccessible while that is essential for timber production, wildlife habitat, and fire management. Traditional multispectral satellite data have some capacity to provide information on forest structural attributes relying on spectral reflectance and vegetation indices. However, dense and complex forest structure tends to saturate traditional broadband vegetation indices, which means that moderate resolution spectral information alone cannot be used to predict vertical forest structure (Hudak et al. 2006; Masek et al. 2015). These multispectral satellite datasets may be insufficient to provide information about forest structure underneath the

canopy (Arroyo et al. 2010). There is scope for the integration of LiDAR and multispectral satellite data to improve the prediction of vertical forest structure. Also, the optimum resolution of satellite datasets for predicting the structural composition of wet eucalypt forests was previously unknown. The study presented in Chapter 5 assessed the effectiveness of integrating various remote sensing datasets to extend the scope of LiDAR-derived variables (used in Chapter 4) and assessed the potential for low-cost data acquisition for large geographical areas.

A key novelty of this research lies in the fusion of the derivatives of three different data sources (Airborne LiDAR, WorldView-3, and Landsat-8 (OLI)) and assessment of the optimum spatial resolution suitable for modelling vertical forest structure (i.e. canopy density) using random forest regression. This research is an extension of Chapter 4, focusing on the problems of integrating LiDAR data with multispectral satellite data for vertical forest structural mapping. Chapter 5 presents a low-cost method for the prediction of the vertical structure of a wet eucalypt forest landscape using vegetation indices, texture features, and topographic attributes. Field data were not used to conduct this research and instead, high-density airborne LiDAR data were applied for training and validation purposes. This research work used thirteen data combination schemes and three pixel sizes to examine the robustness of the forest structural models using a large number of predictor variables. The LiDAR data were processed for noise removal, normalization, derivation of DTMs and quantification of the density of three canopy layers for each of the 1.60 m, 7.5 m and 30 m spatial resolutions. The selected eight texture features and spectral indices were extracted from each band of Landsat-8 (OLI) and WorldView-3 satellite imagery. The high-resolution WorldView-3 data were resampled to the 30 m Landsat-8 resolution for comparison purposes. The high-density discrete return LiDAR point clouds (approximately 30 points/m^2) were thinned to simulate typical operational LiDAR data (about 5 points/m^2) to generate three DTMs of the 1.60 m, 7.5 m and 30 m spatial resolutions to derive topographic attributes separately.

The analysis of this research work showed that the schemes of the 30 m Landsat-8 (OLI) dataset produced the best overall result for all three-pixel sizes with R^2 values of 0.65, 0.46 and 0.44 for the upper, middle and lower layers respectively. Models validated using independent datasets confirmed the robustness of the developed models as the differences between model accuracies and validation accuracies were small (<5.67%) for all the models, indicating the models are transferrable to other areas of the forest landscape. The LiDAR-derived topographic attributes made a substantial

additional contribution to the Landsat dataset for the 30 m pixel size. Surprisingly, WorldView-3 data of 1.6 m pixel size were not a useful predictor for modelling vertical forest structure. Predictive accuracy for modelling vegetation density was highest in the upper canopy layer compared to the middle and lower layers. Texture features and spectral indices separately could predict vegetation density with mean R^2 values of 0.47 and 0.48 respectively for the upper layer. They could predict the vegetation density with mean R^2 values of 0.25 and 0.17 for the lower layer and 0.35 and 0.23 for the middle layer respectively indicating the challenges of predicting the lower and middle layer using multispectral datasets alone. However, mean R^2 values with the combinations of spectral indices and texture features with topographic attributes had some improvement of up to R^2 values of 0.46 for the middle layer and up to 0.44 with the combinations of spectral indices and texture features with topographic attributes and spectral bands of the Landsat-8 dataset for the lower layer. The fused datasets integrating texture features and spectral indices with the LiDAR-derived topographic attributes provided greater R^2 values (0.63 for the upper layer).

Finally, it can be concluded that the fused Landsat-8 (OLI) satellite imagery and the simulated operational LiDAR data provided the overall best result, and so the schemes of the Landsat-8 (OLI) datasets with the resolution of 30 m were reasonably useful for predicting the density of three vertical forest layers. Landsat-8 imagery is freely available high-quality data, so the outputs could be easily adopted by forest managers and planners. Texture features and spectral indices were found to be useful for predicting the upper layer of the wet eucalypt forests. The analysed results showed that WorldView-3 data of 1.6 m resolution are not appropriate to predict vertical forest structure. The analyses in this research work showed that the model accuracies increased with an increasing number of predictor variables indicating the merits of integrating topographic attributes, texture features, and spectral indices.

6.2.4 Comparison of approaches for assessing vertical forest structure

Chapter 4 and Chapter 5 in this research utilised high-density LiDAR data for random forest model training and validation purposes and highlighted the cost-effective approach. Both the chapters confirmed that the larger pixel size (e.g. 30 m Landsat-8) dataset was better suited for predicting vertical forest structure (i.e. canopy density layers); whereas the smaller pixel size datasets (e.g. 1.60 m WorldView-3) were not appropriate for wet eucalypt forests. Chapter 4 tested five pixel sizes to determine the optimal resolution, variable importance and the best parameters (*mtry* and *ntree*) for random forest regression and found the influence of the variable selection on the

random forest results. Chapter 5 examined three pixel sizes that matched the VNIR bands of WorldView-3 (1.60 m), SWIR bands of WorldView-3 (7.5 m) and Landsat-8 (30 m) datasets, and considered texture features and spectral indices with thirteen schemes of datasets for modelling vertical forest structure.

Chapter 4 used topographic attributes derived from high-density LiDAR and a geology vector map to predict understory structure. Chapter 5 extended the approach deploying multispectral satellite data (Landsat-8 and WorldView-3) and illustrated the merits of their derivatives (i.e. texture features and spectral indices) with topographic attributes derived from the simulated operational LiDAR and a geology data to explore the cost-effective and robust modelling approach for large geographical areas. The approach in Chapter 5 was able to confirm that the freely available Landsat data products could be utilised to predict vertical forest structure integrating with DTM derivatives with only a slight sacrifice in the model accuracies if LiDAR data and field inventory data are not available for forest managers and planners.

6.2.5 Contributions to the discipline of remote sensing in forests

This research thesis has presented innovative remote sensing analyses in object-based tree species classification and forest structural mapping using multi-source remote sensing data. The major contributions of this thesis to the current state of knowledge are outlined under three major categories: wet eucalypt forests, object-based image analysis, and applications of remote sensing.

Wet eucalypt forests

- This research thesis utilised the fused airborne LiDAR and hyperspectral remote sensing datasets for tree crown segmentation and species classification. The approach developed in this thesis can be a useful complement to the current aerial photo interpretation mapping technique in wet eucalypt forests in Tasmania.
- The developed approaches to mapping forest species distribution and structure can be useful tools for forest managers and planners for sustainable forest management and biodiversity conservation; e.g. for planning harvest operations and developing conservation strategies.
- This thesis demonstrated that it is possible to predict understory vegetation structure from topographic attributes and geology datasets. This new approach could help predict habitat suitability, bird species richness, fire hazard

prediction and planning for sustainable forest management in wet eucalypt forest ecosystems.

- The research highlighted the importance of particular geology types for driving structural composition in wet eucalypt forests. For example, Permian sediments were found to play a significant role in forest species composition and structure. This information adds to the current state of knowledge to the management of wet eucalypt forests and allows further investigating the influences of geological properties on understory species.

Object-based image analysis

- This research thesis examined the optimum segmentation parameters for vegetation classification in wet eucalypt forest and explored a need for an automated segmentation method to reduce or eliminate the use of a "trial and error" approach. This novel approach used the combinations of different derivatives of airborne LiDAR and hyperspectral data for tree crown segmentation. This is important progress in the field of object-based image analysis.
- This thesis predicted species-specific classes, for example, overstory *Eucalyptus obliqua* and understory forest species *Dicksonia antarctica* based on objects using a random forest classifier. This novel approach utilised the optimized segments from the fused dataset and contributed to the understanding of requirements for data fusion in object-based image analysis.

Applications of remote sensing

- This research thesis utilised high-density airborne LiDAR data for training and validation purposes, and simulated datasets of LiDAR and WorldView-3 for the analyses. These are new developments in the field of remote sensing and reduce the cost and time for field data collection.
- This research utilised a large number of topographic attributes, texture features and spectral indices for predicting the densities of forest canopy layers using different sources of remote sensing datasets. This is the first comprehensive research developing a novel and robust workflow for predicting vertical forest structure of similar types of forests using random forest regression techniques.
- This research showed that the freely available 30 m Landsat-8 (OLI) satellite data had better outcomes than high-resolution (1.6 m) commercial WorldView-3 satellite data for predicting vertical forest structure. This outcome confirms

that the novel approach using moderate spatial resolution multispectral satellite data can be useful for mapping vertical forest structure and helps improve an understanding of the spatial and spectral resolutions required to predict the vertical structure of wet eucalypt forests.

6.2.6 Limitations and future research directions

Although this research thesis has made substantial contributions to the current state of knowledge in the field of mapping wet eucalypt forest species and vertical structure, there are still several considerations that need to be addressed in future research.

Limitations:

- The research in Chapter 3 was conducted using field data from a single 1.6 ha sample plot. Expanding the scope of research to other sites would, therefore, be valuable to further test and validate the usefulness of hyperspectral data for mapping species composition in a variety of wet eucalypt forest settings. This research cannot confirm the best classification approaches in relation to influencing factors, such as forest types, quality of training data, digitized polygons for assessing segmentation accuracies, and classification method used. For example, a lot of the middle layer and lower layer trees were obscured by a relatively dense overstory, so most of the middle layer and lower layer trees and a number of tree species were not included in the training and validation datasets.
- The classification could identify 4-5 classes only whereas the number of tree species is much higher in wet eucalypt forests than those classified in this research.
- The research in Chapters 4 and 5 utilised high-resolution discrete return LiDAR data for training and validation purposes. However, this research had limitations for practical application due to the lack of field data to validate the achieved results. For example, the plant species grown at specific sites cannot be predicted without the detailed information of the upper layer that affects species composition and understory vegetation (Thomas et al. 1999).
- As this research was confined to the predominant mature wet eucalypt forests, this research may not be applicable to other more open forest types and wet forests with younger aged trees that regenerated after timber harvesting or recent wildfires.

Future research directions:

- Since most of the studies have used time-consuming "trial-and-error" segmentation technique to determine the best segmentation depending on subjective visual observation of the segments for object-based classification, this study developed an extensive and systematic approach and recommended research into developing automated algorithms for object-based segmentation, therefore eliminating the "trial-and-error" segmentation approach.
- The future research should replicate the species classification approach presented in Chapter 3 to other types of landscapes and over large geographical areas where detailed field inventory data are available to assess the robustness and transferability of the conceptualized workflow.
- The approaches developed in Chapters 4 and 5 need to be tested over a larger geographical area with a greater diversity of topographical attributes, geology types, texture features, and spectral indices.
- Although the impacts of five key topographic attributes on the understory models were assessed, more research is required into the impacts of individual topographic attributes on the density models of understory structural layers.
- Because of the limited extent of the study landscape, further research could investigate the specific relationships between geology types and vegetation structure. This is because geology and soil nutrients are important elements for forest management and restoration (Wang et al. 2015). Fewer soil nutrients theoretically reduce the growth of the upper layer of vegetation compared to the forest on soils with higher nutrient levels.
- Future research can look into integrating the topographic attributes derived from the recently launched space-borne Global Ecosystem Dynamics Investigation (GEDI) LiDAR data with Landsat-8 (OLI) imagery for further improving the workflow presented in Chapter 5.

6.2.7 Final remarks

The fusion of remote sensing data has become a useful tool to accurately assess and map forest species and vertical structure for planning and monitoring sustainable forest management and biodiversity conservation, with excellent scope to replace the traditional, subjective and time-consuming mapping technique of aerial photo interpretation. This research thesis identified several challenges in remote sensing data

processing and analysis workflows, but the findings and approaches presented in this thesis addressed some of the issues that can improve mapping accuracies and provide guidelines for a robust future investigation. This thesis exemplified that the fused datasets were robust and improved the species classification accuracies where LiDAR-derived CHM had a crucial role. The approach developed in this investigation was able to predict understory structural layers using the selected topographic attributes and mapped geology data. A practicable low-cost and novel approach developed in this study was the addition of spectral indices and texture features to the topographic attributes derived from the simulated LiDAR and WorldView-3 satellite data. In addition, this is state-of-the-art research where high-density discrete return LiDAR data were utilised for training and validation purposes instead of direct field inventory data. This work highlighted the potential of freely available Landsat-8 (OLI) data. This study also demonstrated the utility of machine learning random forest techniques for predicting and mapping species classification and vertical structural layers of a wet eucalypt forest. The mapping approaches of this thesis should be easier and faster than the PI-type mapping approach, and useful for forest managers and planners for sustainable forest management and biodiversity conservation. Thus, this research addresses data complexities, including multidimensionality and nonlinearity in multisource data, and provides a robust approach for the assessment of wet eucalypt forest composition and structure.

Appendix A

Appendix A-1 Forest species classification using LiDAR and hyperspectral data combined or alone in different systems illustrating study area, datasets, purpose, classification algorithms, and accuracies.

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
1	Changshu City in Jiangsu Province in East China and Huanshui Park in the north of Henan Province, Central China	Hyperspectral imagery and LiDAR data	Classify the tree species and evaluate the results	*k*-nearest neighbor (*k*NN), Maximum-likelihood classifier (MLC), Logistic regression and SVM	SVM outperformed all other classification methods ranging from 50% to 67%	(Yang et al. 2019)
2	Toulouse city, France	HySpex visible near infra-red (VNIR) and Short-Wavelength Infra-red (SWIR), panchromatic image and normalized DSM (digital surface model)	Identify the best object-based fusion strategy taking advantage of the complementarity of several heterogeneous airborne data set for 15 tree species	OBIA, and SVM	The overall accuracy of fused data and VNIR data alone are 77% (kappa =74%) and 75% respectively	(Aval et al. 2019)
3	Penobscot and Howland experimental forests in Maine, USA	Hyperspectral imagery and LiDAR data	Classify dominant tree species an compare the different machine learning methods	Decision trees, CN2 rules, K-nearest neighbors, Neural networks, random forest, SVM, Naïve Bayes,	67% for 10 and 59% for 15 dominants species at Howland and Penobscot experimental forests	(Marrs and Ni-Meister 2019)
4	Boreal forest in Norway	Hyperspectral imagery and LiDAR	Compare classification accuracies	Neural network, indices method, partial least	Classification accuracies ranged from 73% to 93%	(Trier et al. 2018)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
	(Gran, Lunner, and Jevnaker municipalities)			squares, deep learning image classification		
5	Qi'ao Island in northwestern Dawei Bay in Zhuhai City, Guangdong Province, China	UAV Hyperspectral imagery and LiDAR data	Classify and map mangrove species	OBIA, *K*NN, and SVM	Overall accuracy of 76.12% (Kappa = 0.73) and 82.39% (Kappa = 0.801) for KNN and SVM respectively	(Cao et al. 2018)
6	Bavarian National Park, South-eastern part of Germany	Hyperspectral imagery and LiDAR data	Discriminate species in classical plant taxonomy	3D segmentation and random forest	Overall accuracy of fused hyperspectral and LiDAR dataset = 83.7%, LiDAR=65.1% and hyperspectral=69.3%	(Shi et al. 2018)
7	Heihe River Basin, China	Hyperspectral imagery and LiDAR data	Crown level tree species classification	Watershed segmentation and classification based on spectral and shape features	Fused LiDAR and hyperspectral dataset (OA = 85.12%, K = 0.90), LiDAR-metrics method (OA = 79.86%, K = 0.81) and hyperspectral-metircs method (OA = 71.26, K = 0.69)	(Wang et al. 2018)
8	Brussels, Belgium	Hyperspectral imagery and LiDAR data	Object-based tree segmentation and monitor tree defoliation and discoloration of broadleaved trees	Partial Least Squares Regression (PLSR	Segmentation accuracy = 91% and accuracies of detection of the healthy tree and unhealthy trees were 93% and 71% respectively	(Degerickx et al. 2018)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
9	Surrey city, British Columbia, Canada	Hyperspectral imagery and LiDAR data	Classify and map 15 common urban tree species	Random forest	Overall accuracy ranged from 51% to 70% and for two native conifer species from 78% to 91%	(Liu et al. 2017)
10	Yushan Forest in Changshu Town in Jiangsu Province, China	Hyperspectral imagery and LiDAR	Delineate and classify tree species	RF	Overall accuracies ranged 85.4-89.3% using whole crown metrics and ranged 87.1-91.5% using sunlit crown metrics	(Shen and Cao 2017)
11	Eastern side of Coromandel region, New Zealand	LiDAR, QuickBird and GIS topography indices	Identify a single native tree species	OBIA, RF, and SVM	Accuracy of fused LiDAR and Spectral data (85.4%) and spectral data alone (75.8%)	(Pham et al. 2016)
12	Southeast Seattle, Washington, and Washington Park Arboretum	Hyperspectral and LiDAR	Test effectiveness of tree species classification	OBIA, RF and multi Class Classifier (MCC)	RF (87.0%) and MCC (88.9%), and overall accuracy = 79.6%	(Zhang et al. 2016)
13	Gulf Islands National Park Reserve, in coastal South-western Canada	Hyperspectral Airborne Imaging Spectrometer and ALS	Map 11 tree species	SVMs	Producer's and user's accuracies for most species ranges >52–95.4% and >63–87.8%, respectively for 11 species	(Jones et al. 2010)
14	The south-eastern part of the Province of Trento, Italy, in the Southern Alps	Hyperspectral image, VHR multispectral imagery, and LiDAR	Understand the level of classification accuracy	SVM and RF	General macro-classes, forest types, and single species, reaching high kappa accuracies	(Dalponte et al. 2012)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
					(93.2%, 82.1%, and 76.5%, respectively).	
15	North of Karlsruhe in the federal state of Baden-Württemberg in Germany	Hyperspectral imagery and LiDAR	Bridge the knowledge gap in understanding the scale effect in imaging (4m to 30m)	SVM and RF	Kappa value= 0.83 for 8m and 0.70 for 30m spatial resolution	(Ghosh et al. 2014)
16	Mediterranean site south of Aix-en-Provence, France	Imaging spectrometry and LiDAR	Map fuel types and properties	SVM	OA = 71.23% (kappa of 0.667)	(Koetz et al. 2008)
17	Finland	Hyperspectral and LiDAR	Test Classification performance	SVM	OA = 67.5%	(Puttonen et al. 2010)
18	The Turtle Creek in north Dallas, Texas	LiDAR and Hyperspectral	Generate a species-level map	Neuro-fuzzy Classification	OA = 68.8 % (Kappa= 0.66)	(Zhang and Qiu 2012)
19	Hawaiian and Panamean tropical rainforests	Hyperspectral and LiDAR	Propose a method for hyperspectral image segmentation	binary partition tree (BPT) algorithm	up to 54.4% for Hawaii and 68% for Panama in the best cases	(Tochon et al. 2015)
20	Complex forest scene of the "Bosco della Fontana" in the Po Plain near the city of Mantua, Italy	Hyperspectral and LiDAR	Analyze the joint effect of hyperspectral and LIDAR data for the classification of complex forest areas	SVMs and Gaussian maximum likelihood (GML) classifier	OA =71.7%	(Dalponte et al. 2008)
21	Tama Forest Science Garden in Tokyo, Japan	Hyperspectral and LiDAR	Propose a methodology for individual tree classification	SVM	OA=82%	(Matsuki et al. 2015)
22	Tama Forest Science Garden, Forestry and Forest Products Research Institute (FFPRI), Japan	ALS and Hyperspectral	Compare two classification approaches which are object-based and pixel-based image analysis	Maximum likelihood method	Pixel-based HSS (Kappa=0.625), pixel-based DCHM combined (Kappa=0.623) and pixel-based DCHM	(Murakami et al. 2011)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
					combined-PC (Kappa=0.620)	
23	Nanawale	Imaging Spectroscopy and LiDAR	Identify and map individuals of nine tree species	semi-supervised SVM	Balanced accuracy (71.1% -73.3%)	(Féret and Asner 2012)
24	Municipality of Aurskog-Høland, southeastern Norway	hyperspectral and ALS data	Estimation of forest attributes	SVM	manual and automatically delineated ITC (individual tree crown) were 0.81 and 0.74, respectively	(Dalponte et al. 2014)
25	Kruger National Park, eastern South Africa	Hyperspectral and LiDAR	Map large-scale species	SVM	OA=76%	(Colgan et al. 2012)
26	Santa Barbara, California, USA	Hyperspectral and LiDAR	Map 29 common tree species	Canonical discriminant analysis	Species-level 83.4% (kappa = 82.6) and leaf-type level with 87.9% accuracy	(Alonzo et al. 2014)
27	Five rainforest sites on the Island of Hawaii, USA	Airborne imaging spectroscopy and LiDAR	Detect invasive species in Hawaiian rainforests	*k*-means clustering algorithm	OA ranged from 63% to 91%.	(Asner et al. 2008)
28	Liangshui National Nature Reserve, Heilongjiang Province, China.	Airborne Hyperspectral and LiDAR Data	Map dominant tree species	SVM	Overall accuracy (86.88%) and Kappa coefficient (0.836) than hyperspectral data only (80.67%, 0.783)	(Liu et al. 2011)
29	Beecroft Peninsula, NSW, Australia	Hyperspectral imagery	Classify Australian forest species at the leaf, canopy and community levels	SVM), AdaBoost and RF	RF (94.7%) and SVM (84.5%)	(Shang and Chisholm 2014)
30	Olney State Forest, the central coast of	Compact Airborne Spectrographic Imager 2 (CASI-2)	Map both spectrally complex species and species groups (subgenus	Supervised maximum	Whole crown and sunlit/ shaded classification	(Goodwin et al. 2005)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
	New South Wales, Australia		grouping) in an Australian eucalypt forest	likelihood classifications	accuracies 80% and 81.7% respectively, overall accuracy=78.9% (Kappa of 0.74)	
31	Kruger National Park, eastern South Africa	Hyperspectral, WorldView-2 and Quickbird data	Map seven common Savannah tree species or genera	Maximum likelihood classifier	OA = 79%±1.8	(Cho et al. 2012)
32	A portion of the Lake Okeechobee watershed in the central Everglades	Hyperspectral and LiDAR	Design framework for vegetation mapping in complex wetlands	RF, SVM, and k-NN	OA = 86% (Kappa value of 0.82)	(Zhang 2014)
33	Pentezug-puszta, part of Hortobágy National Park (Eastern Hungary)	Airborne Hyperspectral imagery	Test the applicability of different image classification methods	Maximum likelihood classifier-MLC), machine learning algorithms (SVM and RF)	Using 30 random training pixels, SVM: 82.06%; RF: 79.14%; MLC: 80.78%, but using 10 random training pixels, SVM: 79.57%, RF: 76.55%, MLC: 52.56%	(Burai et al. 2015)
34	Bavarian Forest National Park (BFNP) at the border between Germany and the Czech	Hyperspectral data	Map bark beetle-induced tree mortality	SVM	84%–96%	(Fassnacht et al. 2014a)
35	Demmin and Karlsruhe, Germany	Hyperspectral data	Compare classification approaches	SVM and RF	86% and 84% respectively for Karlsruhe and Demmin sites	(Fassnacht et al. 2014b)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
36	Tumut region of southern New South Wales, Australia	Hyperspectral data	Performance of MESMA in classifying and mapping nine eucalypt tree species	Multiple endmember spectral mixture analysis (MESMA) maximum likelihood classification	75% (Kappa 0.48) for non-continuum removal (CR) spectra and 83% (Kappa 0.63) for the CR spectra	(Youngentob et al. 2011)
37	Kissimmee River watershed of south-central Florida	Hyperspectral data	Evaluate the applicability of fine spatial resolution data	SVM and RF	90% & 85% for 14 and 55 vegetation communities respectively	(Zhang and Xie 2013)
38	The Florida Keys for benthic habitat mapping	Hyperspectral data	Map benthic habitats	Maximum Likelihood (ML) classifier	84.3% for the group-level & 86.7% for a code-level classification	(Zhang et al. 2013a)
39	Two different landscapes in Wallonia, southern Belgium	Hyperspectral imagery from UAS	Characterize riparian forest health condition	OBIA	79.5% and 84.1 % for five classes for site 1 and 2 respectively	(Michez et al. 2016)
40	The Florida Keys for benthic habitat mapping	aerial photograph and AVIRIS hyperspectral imagery	Design a framework for automated benthic habitat mapping	SVM, RF, and *k*-NN	88.5% and 83.5% for the group-level and code-level classifications, respectively	(Zhang 2015)
41	UNESCO Biosphere Reserve, A temperate Austrian forest, Austria	Hyperspectral imagery	Segment tree crown and classify 13 tree species	Random forest classifier	Overall classification of 13 species = 91.7% (*k*=0.91) based on manual delineation, and 89.4% (*k*= *0*.0.88) based on mean shift segmentation method	(Maschler et al. 2018)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
42	Tihany Peninsula of Lake Balaton, Hungary	Hyperspectral imagery and LiDAR-derived digital canopy model	Map aquatic vegetation	Maximum Likelihood Classifier and SVM	Accuracy using SVM 84% (Eagle data) and 68% (Indices data)	(Stratoulias et al. 2018)
43	Sundarbans (UNESCO World Heritage site) and Bhitarkanika (Ramsar wetland), India	AVIRIS-NG hyperspectral imagery	Classify mangrove species	Spectral Angle Mapper (SAM)	Post-classification accuracy of vegetation = 84.5% (k=0.78)	(Chaube et al. 2019)
44	Yellowstone National Park, the oldest park in the United States	Hyperspectral data	Quantify seedling density in post-fire regeneration sites	Hierarchical classification	Hierarchical Level I classification (78.8%) Level II stand density classification (89.1%)	(Moskal and Jakubauskas 2013)
45	Turtle Creek Corridor, Dallas, Texas, USA	WorldView-2 and LiDAR	Classify with a fusion of the waveform LiDAR and HSR imagery at the object level	Kullback–Leibler divergence-based classification method	OA=97.58% (Kappa=0.97)	(Zhou and Qiu 2015)
46	the Serra de Monchique in southern Portugal	airborne multispectral and ALS	Provide robust indicators of conservation status	Clustering algorithms	70% (11 land cover classes) and 81% (5 classes)	(Simonson et al. 2012)
47	southwest of the city of Jyväskylä, Central Finland	LiDAR and WorldView-2	Develop a classification workflow for boreal forest habitat type mapping	RF and OBIA	Overall accuracy 78.0% (ranges from 73% to 79%)	(Räsänen et al. 2014)
48	eastern Connecticut, north-eastern USA	LiDAR and aerial orthophotographs	Propose a fully automated rule-based algorithm to develop a 1 m resolution land cover classification	GEOBIA rule-based algorithm	Overall accuracy for the 8-category classification was 93.1% (kappa = 0.90), But 5-class land cover map was 94.8% (kappa = 0.92)	(Parent et al. 2015)
49	eastern Strzelecki Ranges, southeast	LiDAR data and WorldView-2	Test the effectiveness of the different	OBIA with decision	82.35% after fusing all eight bands and the	(Zhang et al. 2014)

S.N.	Study area	Dataset/s	Purpose	classification algorithms	Accuracy	Reference
	Victoria, Australia		input data layers	trees in CART 6.0 software	LiDAR data, but LiDAR data alone (61.39%) or four conventional bands only (61.42%)	
50	Heiberg Memorial Forest and adjacent State Forest lands, Central New York State	LiDAR and QuckBird imagery	Synergistic use of QuickBird and LiDAR data for species classification	OBIA and decision tree classification	Highest classification accuracy, Kappa = 91.6%	(Ke et al. 2010)
51	meander areas on the Garonne and Allier Rivers, France	LiDAR	Classify forest and ground types	Unimodal and bimodal distribution skewness and kurtosis models	66 and 98% for Five types of riparian forest	(Antonarakis et al. 2008)
52	Southwestern Germany close to the city of Karlsruhe	LiDAR	Investigate a comprehensive set of variables derived from secondary LiDAR parameters concerning their tree species discrimination	Gaussian scatters, linear discriminant analysis and statistical values	The overall accuracy of 57% classified for six tree species, 78% for four species	(Heinzel and Koch 2011)

Appendix A-2 Forest structural attributes using LiDAR and hyperspectral and their derivatives combined or alone in different systems illustrating RS datasets, key structural attributes, the purpose of study, algorithms, and accuracies.

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
1	VHR, Hyperspectral, and LiDAR	CHM, Object texture and geometry features	Develop a framework for mapping crop species	Heihe River Basin, Gansu Province, China	SVM	OA = 90.33% (kappa of 0.89)	(Liu and Bo 2015)
2	Hyperspectral and LiDAR	Maximum LiDAR height mean LiDAR height, the standard deviation of LiDAR height. Coefficient of variation of Light LiDAR height, percentiles of LiDAR height, Laser intercept index	Estimate forest below-and above-ground biomass, and total biomass	Heihe River Basin in Zhangye City of Gansu Province northwest China	Partial least squares (PLS)	LiDAR metrics (R^2 = 0.74), CASI-metrics (R^2 = 0.51), LiDAR and CASI-metrics (R^2 = 0.78)	(Luo et al. 2017)
3	Hyperspectral and LiDAR	Quantile height, mean height, the standard deviation of heights, skewness of heights	Develop a method for mapping forest canopy species diversity	Shennongjia National Forest Natural Reserve in Xingshan country, Hubei province, China	RF	Height accuracy (R^2 = 0.90) Individual tree crown isolation (R^2 = 0.87) Shannon-Wiener biodiversity index (R^2 = 0.83)	(Zhao et al. 2018)
4	Hyperspectral and LiDAR	Species, CHM, maximum crown height, and crown area, an abundance of each species, LAI	Large-scale species mapping	Kruger National Park, eastern South Africa	SVM	76% achieved for 15 species	(Colgan et al. 2012)
5	Hyperspectral and LiDAR	DEM, DSM, CHM, canopy form(Height and size), Species	Propose a methodology to	Tama Forest Science Garden, Tokyo, Japan	SVM	OA = 68% (48% for hyperspectral data only)	(Matsuki et al. 2013).

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
			classify tree species				
6	Hyperspectral and LiDAR	Species, CHM, intensity variables	Identify and map nine tree species by comparing a variety of semi-supervised classifiers	Hawaiian lowland	SVM	80%-100% for eight species	(Féret and Asner 2012)
7	Hyperspectral and LiDAR	species diversity, individual tree metrics	Identify urban tree species at the individual tree level	Turtle Creek in north Dallas, Texas	Neuro-fuzzy Classification and defuzzification	The overall accuracy of 86% and a Kappa value of 0.82	(Zhang and Qiu 2012)
8	Hyperspectral, GeoEye-1, and LiDAR	Percentiles (5%, 10%, 25%, 50%, 75%, 90%, and 95%) Standard deviation, skewness, Kurtosis, Variance, heights (min, max, range, mean) of all height points within each pixel, species	Understand what level of classification accuracy and implications of a downgrading of the data characteristics	The south-eastern part of the Province of Trento, Italy, in the Southern Alps	Gaussian maximum likelihood algorithm, sequential forward floating selection (SFFS) algorithm	Overall accuracy = 88.0%	(Dalponte et al. 2012)
9	AISA hyperspectral and LiDAR	LiDAR-derived height and volumetric canopy profile data	Assess the utility of AISA and map 11 tree species	Gulf Islands National Park Reserve, in coastal South-western Canada	SVMs, McNemar's tests for accuracy	Increased accuracies both producer's (+5.1–11.6%), user's (+8.4–18.8%)	(Jones et al. 2010)
10	Hyperspectral and LiDAR	species, elevation, and intensity	Analyze the seasonal effect on differentiating tree species in an urban environment	University of Northern Iowa campus (49 ha), Cedar Falls, Iowa, USA	hierarchy (nearest neighbour) classification	classification improved the results by 19%	(Voss & Sugumaran, 2008

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
11	AVIRIS	leaf-type, leaf/plant duration and life form, leaf-type, leaf/plant duration, and life form	Impact of spatial resolution on the classification of plant species and functional	Smithsonian Environmental Research Center, Wind River Experimental Forest, and Sierra National Forest	Canonical Discriminant Analysis	Overall accuracy ranging from 61 to 96%	(Roth et al. 2015)
12	Hyperspectral and LiDAR	Shape-based classification parameters, species classification	Classify species using parameters	FARO, Lake Mary, USA	SVM	Overall accuracy = 67.5% for all trees	(Puttonen et al. 2010).
13	AVIRIS imagery and LiDAR	Stem count, Canopy area, species, crown widths at selected heights, ratios of crown heights to widths, distributions of intensity through the crown, crown porosity	Delineate individual crowns from a gridded canopy maxima model, USA	Santa Barbara, California	Canonical varieties and classification	83.4% (kappa = 82.6) and 93.5% for Species-level and leaf-type level respectively	(Alonzo et al. 2014)
14	Compact Airborne Spectrographic Imager (CASI) and LiDAR	Five-class set (Grasses, herbs, bare, pioneer communities, low woody vegetation, bushes	Classify the floodplain vegetation, and compare the classification results	Waal River, one of the main branches of the river Rhine in the Netherlands	maximum-likelihood classifier, neural network and CART decision tree classifiers	OA = 81% (five-class set), but Using 74% for CASI data only	(Geerling et al. 2007)
15	QuickBird, AISA hyperspectral and LiDAR	Elevation and intensity, tree species, heights	Develop a methodology for the identification of tree species in an urban environment	University of Northern Iowa campus (121 acres) in Cedar Falls, Iowa	Hierarchy (nearest neighbour rules) classification	In 2006, 81-93% (kappa, 71-90%), in 2004 images, 92% (Kappa, 0.88) and in 2003 QuickBird, 88% (Kappa of 0.82)	(Sugumaran and Voss 2007)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
16	Hyperspectral and LiDAR	Tree species, species diversity, species richness	Modelling vascular plant diversity patterns	Bavarian Forest national park in Germany	Boruta algorithm, random forests	R^2_{LiD} = 0.75, R^2_{MNF} = 0.76, $R^2_{MNF+LiD}$ = 0.78	(Leutner et al. 2012)
17	Satellite hyperspectral and LiDAR	Mean and median digital crown model, Shannon index, species richness, Vegetation Indices and ratios	To generate a spatial prediction of vascular plant richness	Andes foothills of the Maule Region, Chile	RF algorithm, multiple linear regression, Spatial autocorrelation	Adjusted R^2 = 0.571, RMSE = 5.05	(Ceballos et al. 2015)
18	Imaging spectroscopy and LiDAR	Tree species (excluding elevation, slope, and other topographical information)	Test our ability to identify patterns in plant species composition	tropical broadleaf rainforest, central Panama	Principal component analysis to reduce the 214 bands	R^2= 0.88, second-order polynomial elevation and slope explained 75% and 59% of the variation	(Higgins et al. 2014)
19	Airborne hyperspectral data (both 4 and 8 m resolution), Hyperion and LiDAR	Selected components of Minimum Noise Fraction (MNF), Vegetation Indices (VI), heights, tree species, reflectance value of all bands, canopy layers. Normalized digital surface model (nDSM)	Produce accurate tree species maps and to examine the influence of spatial resolution on the derived maps	North of Karlsruhe in the federal state of Baden-Württemberg, Germany	SVM and RF	OA = 0.62 -0.86, (kappa = 0.54-0.83)	(Ghosh et al. 2014)
20	Full waveform ALS at night & Landsat ETM	ground elevation, canopy height, the height of median energy, height at 25% and 75% of the cumulative waveform energy, vertical distribution, canopy density, NDVI	Bird richness and abundance	Patuxent National Wildlife Refuge, Maryland, USA.	Automated algorithm and normalized ratio	45% of the variation in bird species richness, and other alpha species diversity predictions (R^2 = 50–60%)	(Goetz et al. 2007)
21	LiDAR	Species richness, Simpson index, tree height (Max, mean), canopy height (max, min,	Model the activity, richness, and composition of	Bavarian Forest National Park in	Canonical correlation analysis	R^2 of up to 0·6	(Müller and Brandl 2009)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
		mean, tree species, herb layer (total cover, height, number of species, deadwood	assemblages of forest-dwelling beetles	South-eastern Germany			
22	LiDAR	Different height metrics, Percentiles for 0, 5%,.. ., 100% of the laser height distributions, canopy density, canopy heights, proportions of laser hits, quartiles of the height distribution, number and Lorey's mean height of understory trees	To characterise the vertical stand structure of heterogeneous boreal forests, prediction of height distributions of small trees for mapping	Kalkkinen, southern Finland	Histogram thresholding method (HistMod), regression models, Lloyd's algorithm	R^2 of 0.76 and standard error is 0.130	(Maltamo et al. 2005)
23	LiDAR and Aerial photos	Tree height, canopy height, canopy cover, stand density, spatial heterogeneity (by PCA), canopy gap volume and size, the spatial distribution of canopy trees (or overstory), Crown diameter, Crown depth, No. of trees, Stem density, Lorey's height, Shannon's Height Diversity Index (SHDI)	Develop a quantitative method based on a novel application of two well-established statistical techniques – lacunarity analysis and PCA to support a broad range of practical applications in SFM, long-term ecological monitoring, and forest science	North America	'Gliding-box' algorithm to compute lacunarity at discrete measurement scales, 3-D crown modelling techniques	Canopy cover (R^2 = 0.85) and gap volume (R^2 = 0.84), attributes of stand and canopy structure, $0.27 \leq R^2 \leq 0.58$)	(Frazer et al. 2005)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
24	Airborne LiDAR	Tree height, Vertical layers: ground, low vegetation (0-1m), medium vegetation (1-5m) and high vegetation, gaps, canopy cover and its vertical density, LAI	Characterize the ecological structure of a forest landscape	Rubicon catchment in the Cradle Coast Region of Tasmania, Australia	Regression analysis	R^2= 0.52; RMSE = 0.91	(Mura et al. 2015)
25	Airborne LiDAR, SPOT 4 and SPOT 5	Maximum vegetation height (mean, SD), plant species, NDVI, Shannon's diversity index for height, percentiles of height above the ground	Predict the stand-scale abundance and species richness of birds and beetles in a managed boreal forest landscape	Middle boreal zone of northern Sweden	Multiple regression models, *k*NN	Stand-scale biodiversity patterns in beetles (R^2=0.47-0.59), and in birds (R^2=0.41-0.42)	(Lindberg et al. 2015)
26	Airborne LiDAR	CHM, mean height, canopy height percentiles (H5 to H95), mean overstory canopy height, canopy height (mean, median, maximum, and standard deviation), canopy cover, Percentage of returns above height threshold of 0.5, 2, 2-8, >8 m)	Modelling organism-habitat relationships focusing on great tit (Parus major) habitat, the relationship between breeding success and forest structure	Monks Wood in Cambridgeshire, eastern England	Organism-Habitat Relationships	Using all returns (R^2=0.674, p=0.002; H_{mean}: R^2=0.661, p=0.002), and using first returns only (R^2=0.713, p <0.001; H_{mean}: R^2=0.718, p <0.001)	(Hill and Hinsley 2015)
27	Very high-resolution satellite imagery and LiDAR data	Tree species, Number of trees, height, Number of snags, species richness, nDSM, Spectral heterogeneity of pixel values, canopy height, Shape index	Focus on Structure (RFI S), for the assessment of habitat quality and Identification of 'hot-spot' areas	Semi-natural Natura 2000 site Salzachauen, Austria	Descriptive statistics, Spearman's rank correlation coefficient, PCA,	expert-based weighting: ρ = 0.066; statistical weighting: ρ = 0.191	(Riedler et al. 2015)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
	calculated on patch level				statistical correlation		
28	Airborne LiDAR	live, dead, and total BA, tree density, crown length (mean, median), Lorey's height (BA weighted mean height), canopy cover, Density, trees/ha	Demonstrate a novel three-step approach to evaluate habitat based on multiple habitat attributes and their influence at multiple grain sizes	Savannah River Site (SRS) in South Carolina, USA	Non-linear seemingly unrelated regression and multiple linear regression	R^2 ranges from 0.34 to 0.81	(Garabedian et al. 2014)
29	Airborne LiDAR and Aerial images	Height (mean & SD), stem number (mean, SD), volume (mean & SD), canopy height, crown segments, species	Develop a nonparametric Semi-individual tree crown approach	Aurskog-Holand, in southeast Norway	area-based algorithms	Relative RMSE (88.28%) and RF, (82.54%); Relative bias (0.98) and, (RF, 8.61%)	(Breidenbach et al. 2010)
30	LiDAR data and large scale (1:4000) photography (LSP)	Plot level foliage/branch projected cover, dominant and co-dominant species, stand and tree heights, Foliage cover, Species distributions	Quantify the floristics and structure of mixed species forests, height and FPC distributions	Near Injune, central east Queensland	Based on LSP interpretation	79% for dominant species; individual tree (R^2= 0.91, S.E. = 1.34 m,) and stand height (R^2 = 0.84, S.E. = 2.07 m)	(Tickle et al. 2006)
31	Small footprint discrete return LiDAR	CHM, Stem count, vegetation cover, bare ground, Canopy Depth, Crown cover, plot-based stem density, Ground and canopy height surfaces	Quantify the carbon and diversity values of forests for national forest monitoring initiatives	Open forests and woodlands near Injune in Central Queensland, Australia	Height-scaled Crown Openness Index (HSCOI)	70–80%	(Lee and Lucas 2007)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
32	Airborne LiDAR	Vertical profile of the canopy, canopy density, canopy height, plant species composition (body size, species, and number of birds excluded from this study)	Predict the composition of bird communities from the plant's composition and structure	Bavarian Forest National Park in southeast Germany	Co-correspondence analysis, statistical modelling	R^2 of up to 0·6	(Müller et al. 2010)
33	Airborne LiDAR	Vegetation height and cover, beetle abundance and species richness	Evaluate the potential of ALS and forest estimates derived from satellite images	Middle boreal zone of northern Sweden	Multiple regression models, kNN	Biodiversity patterns in beetles (R^2 = 0.47–0.59) and lesser extent in birds (R^2 = 0.41–0.42).	(Lindberg et al. 2015)
34	Airborne LiDAR	Eucalypt top height, basal area, and stems per hectare, top height, height percentiles, intensity, canopy structure	Investigate how metrics derived from laser scanning data	Central Forest Management Area in Victoria, Australia	Regression modelling, Exploratory data analysis	Top height (R^2 = 0.87; root mean square error (RMSE) = 3.9 m) using a single height percentile variable	(Haywood and Stone 2011)
35	Airborne LiDAR	Canopy height, basal area, and LAI, mean vegetation height, canopy cover, height percentiles, 1st, 2nd, 3rd, returns, % of vegetation returns, heights (min, max, range, mean, median, modal, standard deviation, CV)	Evaluate the use of lidar data for characterizing forest successional stages	Palouse Range, Clearwater Mountains in Northern Idaho, USA	Random Forests, the rule-based classification procedure	Overall accuracy = 95% for six species	(Falkowski et al. 2009)
36	Airborne LiDAR	CHM, crown sizes, number of trees, species	Propose a framework for the simplicity of a	Forest site near Sault Site. Marie, Ontario, Canada	Watershed segmentation, semivariogram, multi-scale ITC	OA=74% (plot with mixed wood) and 72%	(Hu et al. 2014)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
			CHM-oriented method		delineation method	(plot with deciduous trees)	
37	Airborne LiDAR	Species, richness, AGB, canopy strata, canopy density, and elevation metrics	Accuracy of predictions of above-ground biomass and plant species richness of tropical dry forests	Private reserve in the centre of the Yucatan Peninsula	multiple regression analysis and ANOVA	species richness with RMSE = 5.5 (leaf-on) vs 5.8 (leaf-off) species	(Hernández-Stefanoni et al. 2015)
38	Airborne LiDAR	CHM, DBH, crown base, canopy length, canopy type, canopy heterogeneity, canopy layering	Develop a transferable, grid-based assessment of canopy layering	Temperate forest (800 ha) located in Canton of Aargau, Switzerland	Multi-scale method	>77% for $canopy_{type}$ and $canopy_{heterogeneity}$, and >62% for $canopy_{layer}$ and $canopy_{length}$	(Leiterer et al. 2015b)
39	Airborne LiDAR, SAR/InSAR, Landsat ETM+, QuickBird	Min, max, mean, standard deviation for canopy height, canopy cover and energy of waveforms	Map forest structure	Sierra Nevada mountains of California	Regression models	Mean canopy height ((r^2= 0.61, RMSE = 7.3m), combined (r^2= 0.72, RMSE = 6.4m),	(Hyde et al. 2005)
40	Airborne LiDAR	Species richness, canopy height	Explore global links between tree canopy height and species richness of amphibians, birds, and mammals	Global	Generalized additive models and geographically weighted regressions.	R^2 for birds, mammals, and amphibians are 90.3%, 88.5%, and 87.1% respectively	(Roll et al. 2015)
41	Airborne LiDAR and Worldview-2	CHM, height profiles, vegetation height, land cover classes	to assess the EODHaM (EO Data for Habitat Mapping)	Natura 2000 site the Veluwe, Province Gelderland Province, largest	Regression and correlation	adjusted R^2 = 0.95 for vegetation height, R^2 = 0.96 overall land cover	(Mücher et al. 2015)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
			classification results	end moraine, the Netherlands		classes' accuracy= 74%	
42	Airborne LiDAR	Canopy-structure types, canopy layer, canopy length, species	Propose a method to provide quantitative canopy-structure descriptors on different scales	Laegeren site (200 ha), semi-natural mixed deciduous mountain forest, northwest of Zurich	the hierarchical, multi-scale classification approach	52% to 82% user accuracy (canopy layering) and 89-99% user accuracy (canopy type)	(Leiterer et al. 2015a)
43	Airborne LiDAR	Canopy height, midstory height, midstory density, canopy density, species richness lower canopy height, understory	To explore new approaches to map biodiversity over large areas by modelling species richness of forest songbirds	Baraboo Hills, southern Wisconsin, USA	Correlation coefficients	15 and 20% of the variability in richness within deciduous forest songbird communities	(Lesak et al. 2011)
44	Airborne LiDAR	Height, density, the proportion of echoes above five meters, canopy layer, shrub layer	Examine the habitat selection patterns of moose within a year	West coast of Finland	Linear mixed effects' models modeling (LME)	During autumn, 70–80% of the moose feeding occurred in older forests	(Melin et al. 2016)
45	Airborne LiDAR	Species heterogeneity, canopy height classes, stand profiles, height, and density percentiles	Evaluate the value of LiDAR data for large-scale habitat assessments	Three biogeographic regions within Switzerland: the Swiss Jura, the Northern Pre-alps and the Eastern Central Alps	Boosted regression trees (BRTs),	Explained deviance (D2) = 56.5%, area under the receiver operating characteristics curve (AUC) = 0.98 for the combined model and D2 ranges from 17.2 to 23.9%	(Zellweger et al. 2014)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
						for three models (field, LiDAR and combined)	
46	AVIRIS and Field Diversity Data	woody vascular plant species richness, leaf chlorophyll, water, and nitrogen content, specific leaf area	Biodiversity prediction equation to map and monitor aspects of biodiversity	Hawaiian Lowland Rainforests, USA	Monte-Carlo simulation technique, the diversity prediction algorithm	$R^2 = 0.85$, $p < 0.01$	(Carlson et al. 2007)
47	Airborne hyperspectral sensing, ASTER	Shannon-Wiener biodiversity indices, canopy tree diversity, NDVI, species, reflectance (mean, SD, min, max), vegetation indices	Monitoring forest biodiversity	Sierra Leonean national park, West Africa	Random Forests regression	pseudo-R^2=84.9%	(Vaglio Laurin et al. 2014)
48	Photo interpretation, hyperspectral and the multispectral data	Heights (Min, Max, range, mean, median, SD), 5th, 10th, 25th, 50th, 75th, 90th, 95th percentiles of all height points within each pixel	Classification accuracy, differences between high- and low-point density LiDAR acquisitions	"Val di Sella", south-eastern part of the Province of Trento, Italy, in the Southern Alps	SVM and RF	85.8% at macro class level; hyperspectral for all the bands (SVM OA, 74.1%, Ka of 66.3% and AA of 70.3%) & (RF OA, 70.8%, KA of 61.9% & AA 68.1%)	(Dalponte et al. 2012)
49	Hyperspectral	Forest biophysical properties (e.g. species composition, canopy closure, LAI), forest biochemical properties (e.g. (chlorophyll a and –b), nitrogen, lignin, water	Map forest species and determine biophysical and biochemical properties	Greater Victoria Watershed District, Canada	Canopy radiative transfer model	94.2% from 70% of the training data and 90.0% from the test data	(Goodenough et al. 2006)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
50	Hyperspectral	Shape, absorption depth, crossover, community, canopy, and leaf-level spectral reflectance and species	Demonstrate the species discrimination and investigate the key wavelengths	Beecroft Peninsula, North of Jervis Bay, adjacent to Jervis Bay National Park, Australia	SVM, RF, Adaboost	Leaf level (94.7%), canopy level (84.5%) and community level (75.5%)	(Shang 2013)
51	AVIRIS	Species, LAI, canopy cover, canopy architecture, leaf volume, biochemicals (chlorophylls, carotenoids, anthocyanins, N and canopy water contents)	Explore the contributions of a microsite, substrate, and climate controls on canopy reflectance, and analyse sources of spectral variation	Lowland Hawaiian rainforest ecosystems on the island of Hawaii, eastern Amazon site	Regression analysis	Regression coefficients r =0.35–0.54 for pigments, and r =0.17-0.22 for nutrients	(Asner 2008)
52	Hyperspectral	Ecotope classes, number of trees	Investigate the possibility of using airborne hyperspectral imagery for the classification of ecotopes	Belgium, locally known as the Biological Valuation Map (BVM)	RF, Adaboost and neural network classifier	RF (68.8%), Adaboost (69.5%) and neural network classifier (63.7%), but overall accuracy (close to 70%)	(Chan and Paelinckx 2008)
53	Airborne Imaging Spectroscopy	Species, non-flowering crowns	Identify individuals of three focal canopy tree species amongst a diverse background of tree and liana species	Barro Colorado Island in Gatun Lake, Panama	SVM	producer's accuracies of 94-97% and user's accuracies of 94-100%	(Baldeck et al. 2015)
54	AVIRIS Hyperspectral	Land cover classes	Investigate the performance of	Hekla Volcano in Iceland	SVM and RF	Approx. 90%.	(Waske et al. 2009)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
			SVM and RF comparing with maximum likelihood classifier and spectral angle mapper				
55	Hyperspectral	tissue (leaf and bark), pixel, crown scales, vegetation chemistry and structure, tree species from full-range, Vegetation Pigments, stress, Water Content, Lignin, Cellulose, Nitrogen	Explore a method to classify seven tropical rainforest tree species from full-range (400–2,500 nm)	La Selva Biological Station, Sarapiquí Canton, Costa Rica	Spectral Mixture Analysis	The overall accuracy of 86.8% for leaves, 74.2% for bark, and 84.9% for leaves plus bark	(Clark and Roberts 2012)
56	Aerial photography, CASI, and hyperspectral data	Tree species discrimination, Species distributions, CASI tree crown delineation, crowns	Develop an approach for the generation of tree species maps at the tree crown level	Southern Brigalow Belt, southeast central Queensland, Australia	Stepwise discriminant analysis	The overall accuracy of 87% and 76% for training and testing datasets; n=398	(Lucas et al. 2008)
57	LiDAR and QuickBird	canopy height, Aboveground biomass (AGB) and volume	Introduce the GEOBIA framework to estimate canopy height, above-ground biomass, and volume	Training and Research Forest of Lake Duparquet (TRFLD), Quebec, Canada	SVM and semi-automatic GEOBIA approach	R=0.85 for canopy height, AGB and volume, but RMSE=3.37m, 39.48mg/ha and 52.59m^3/ha respectively	(Chen et al. 2012)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
58	ALS and IKONOS data	Foliage height diversity, LiDAR-derived indices, habitat types, vegetation density index, shrub density indices	Assess the utility of LiDAR data in quantifying vegetation structural characteristics that relate to avian diversity, density, and occurrence	Black Hills Experimental Forest in South Dakota, USA	Shannon–Wiener information index, correlations	shrub density index (r=0.632 at 50 m and r=0.578 at 100 m), tree vegetation index (r=0.624, pb0.001, n=51 for FHD height category D at a 50 m radial buffer)	(Clawges et al. 2008)
59	LiDAR and QuickBird	CHM, Height classes, forest types	Develop an airborne LiDAR sampling strategy to model full-scene forest canopy height from optical imagery, LiDAR transects and GEOBIA	Southwest of Campbell River on Vancouver Island, British Columbia, Canada	multiple regression models and GEOBIA	the combined error of 6.2 m, but 6.0 m and 6.8 m for conifer and deciduous forests separately (Overall accuracy=72.68% and Kappa=0.58)	(Chen and Hay 2011)
60	ALS and Landsat ETM	Canopy height, species distribution, surface topography	Quantify the habitat heterogeneity of a contiguous northern hardwoods forest	Hubbard Brook Experimental Forest, Woodstock, New Hampshire, USA	random forests	the model predicted 39% for 3-year model and >90% for 4 year	(Goetz et al. 2010)
61	LiDAR, Uninhabited Aerial Vehicle Synthetic	CHM, canopy cover, crown diameters, crown area, a product of height and crown diameter	Evaluate the efficacies of different RS data sets in mapping	Unfragmented temperate deciduous forests at White Mountains of New	RF	For all species, >30% (Radar metrics), 35-65% (Landsat metrics) and >50%	(Swatantran et al. 2012)

SN	RS Datasets	Key forest structural attributes	Purpose	Study area	Algorithms	Accuracy	Reference
	Aperture Radar (UAVSAR) and Landsat ETM+		bird prevalence across a forested landscape	Hampshire, USA of New Hampshire, USA		ranging from 54-71% (LiDAR metrics)	
62	Imaging spectroscopy and ALS	Canopy Gap Profile, LAI, canopy structure, canopy composition, relative abundance	Provide spatially comprehensive estimates of forest canopy composition and structure	Prospect Hill region of Harvard Forest LTER (mixed temperate forest in Central Massachusetts)	Terrestrial biosphere model, Ecosystem Demography (ED2) model	R^2 of 0.672 & RMSE of approx. 10m^2/ha; RMSE ranges 85%–104% to 37%-57% for carbon flux predictions	(Antonarakis et al. 2014)

References

Abdel-Rahman, E.M., Ahmed, F.B., & Ismail, R. (2013). Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. *International Journal of Remote Sensing, 34*, 712-728

Ahmed, O.S., Franklin, S.E., Wulder, M.A., & White, J.C. (2015). Characterizing stand-level forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the random forest algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing, 101*, 89-101

Alam, M.M., Strandgard, M.N., Brown, M.W., & Fox, J.C. (2012). Improving the productivity of mechanised harvesting systems using remote sensing. *Australian Forestry, 75*, 238-245

Alexander, D.L.J., Tropsha, A., & Winkler, D.A. (2015). Beware of R^2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of Chemical Information and Modeling, 55*, 1316-1322

Alonzo, M., Bookhagen, B., & Roberts, D.A. (2014). Urban tree species mapping using hyperspectral and LiDAR data fusion. *Remote Sensing of Environment, 148*, 70-83

Amiri, N., Polewski, P., Yao, W., Krzystek, P., & Skidmore, A.K. (2017). Detection of single tree stems in forested areas from high density ALS point clouds using 3D shape descriptors. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W4*, 35-42

Anderson, E.S., Thompson, J.A., Crouse, D.A., & Austin, R.E. (2006). Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM. *Geoderma, 132*, 406-415

Antonarakis, A.S., Munger, J.W., & Moorcroft, P.R. (2014). Imaging spectroscopy- and LiDAR-derived estimates of canopy composition and structure to improve predictions of forest carbon fluxes and ecosystem dynamics. *Geophysical Research Letters, 41*, 2535-2542

Antonarakis, A.S., Richards, K.S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment, 112*, 2988-2998

Armston, J.D. (2013). Assessment of airborne LiDAR for measuring the structure of forests and woodlands in Queensland, Australia. In, *School of Geography, Planning & Environment Management* (p. 197). Australia: PhD Thesis. School of Geography, Planning & Environment Management, The University of Queensland

Arroyo, L.A., Johansen, K., Armston, J., & Phinn, S. (2010). Integration of LiDAR and QuickBird imagery for mapping riparian biophysical parameters and land cover types in Australian tropical savannas. *Forest Ecology and Management, 259*, 598-606

Arvor, D., Durieux, L., Andrés, S., & Laporte, M.A. (2013). Advances in Geographic Object-Based Image Analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing, 82*, 125-137

Asner, G. (2008). Hyperspectral remote sensing of canopy chemistry, physiology, and biodiversity in tropical rainforests. In M. Kalacska, & G.A. Sanchez-Azofeifa (Eds.), *Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests* (pp. 261-296): CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

Asner, G.P., Kellner, J.R., Kennedy-Bowdoin, T., Knapp, D.E., Anderson, C., & Martin, R.E. (2013). Forest canopy gap distributions in the southern Peruvian Amazon. *PLoS ONE, 8*, e60875

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J., & Hughes, R.F. (2008). Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment, 112*, 1942-1955

Attiwill, P.M., Ryan, M.F., Burrows, N., Cheney, N.P., McCaw, L., Neyland, M., & Read, S. (2014). Timber harvesting does not increase fire risk and severity in wet eucalypt forests of Southern Australia. *Conservation Letters, 7*, 341-354

Aval, J., Fabre, S., Zenou, E., Sheeren, D., Fauvel, M., & Briottet, X. (2019). Object-based fusion for urban tree species classification from hyperspectral, panchromatic and nDSM data. *International Journal of Remote Sensing, 40*, 5339-5365

Avitabile, V., Baccini, A., Friedl, M.A., & Schmullius, C. (2012). Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sensing of Environment, 117*, 366-380

Avitabile, V., Herold, M., Henry, M., & Schmullius, C. (2011). Mapping biomass with remote sensing: A comparison of methods for the case study of Uganda. *Carbon Balance Manag, 6*, 1-14

Azaele, S., Cornell, S.J., & Kunin, W.E. (2012). Downscaling species occupancy from coarse spatial scales. *Ecological Applications, 22*, 1004-1014

Baker, S.C. (2006). Ecology and conservation of ground-dwelling beetles in managed wet eucalypt forest: edge and riparian effects. In, *School of Zoology* (p. 310). Private Bag 05, Hobart, Tasmania 7001, Australia: PhD Thesis. University of Tasmania

Baker, S.C., Garandel, M., Deltombe, M., & Neyland, M.G. (2013). Factors influencing initial vascular plant seedling composition following either aggregated retention harvesting and regeneration burning or burning of unharvested forest. *Forest Ecology and Management, 306*, 192-201

Baker, S.C., Grove, S.J., Wardlaw, T.J., McElwee, D.J., Neyland, M.G., Scott, R.E., & Read, S.M. (2017). Monitoring the implementation of variable retention silviculture in wet eucalypt forest: A key element of successful adaptive management. *Forest Ecology and Management, 394*, 27-41

Baker, S.C., & Read, S.M. (2011). Variable retension silviculture in Tasmania's wet forests: ecological rationale, adaptive management and synthesis of biodiversity benefits. *Australian Forestry, 74*, 218-232

Baldeck, C.A., Asner, G.P., Martin, R.E., Anderson, C.B., Knapp, D.E., Kellner, J.R., & Wright, S.J. (2015). Operational tree species mapping in a diverse tropical forest with airborne imaging spectroscopy. *PLoS ONE, 10 (7)*, e0118403

Ballanti, L., Blesius, L., Hines, E., & Kruse, B. (2016). Tree species classification using hyperspectral imagery: A comparison of two classifiers. *Remote Sensing, 8*, 445

Balmer, J. (2016). Floristic response to landscape context in vascular plant communities in Eucalyptus obliqua and Eucalyptus regnans wet forest, southern Tasmania. In, *Discipline of Geography and Spatial Science, School of Land and Food* (p. 501). Australia: PhD Thesis. Discipline of Geography and Spatial Science, School of Land and Food, University of Tasmania, Hobart

Bar-Massada, A., & Wood, E.M. (2014). The richness-heterogeneity relationship differs between heterogeneity measures within and among habitats. *Ecography, 37*, 528-535

Barbier, S., Gosselin, F., & Balandier, P. (2008). Influence of tree species on understory vegetation diversity and mechanisms involved—A critical review for temperate and boreal forests. *Forest Ecology and Management, 254*, 1-15

Barkman, J.J. (1988). A new method to determine some characters of vegetation structure. *Vegetatio, 78*, 81-90

Bartels, S.F., & Chen, H.Y.H. (2013). Interactions between overstorey and understorey vegetation along an overstorey compositional gradient. *Journal of Vegetation Science, 24*, 543-552

Batini, C., Blaschke, T., Lang, S., Albrecht, F., Abdulmuttalib, H., Barsi, A., Szabó, G., & Kugler, Z. (2017). Data quality in remote sensing. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W7*, 447-453

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31

Bergen, K.M., Goetz, S.J., Dubayah, R.O., Henebry, G.M., Hunsaker, C.T., Imhoff, M.L., Nelson, R.F., Parker, G.G., & Radeloff, V.C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for LiDAR and radar spaceborne missions. *Journal of Geophysical Research: Biogeosciences, 114*, 13

Bernstein, L.S., Jin, X., Gregor, B., & Adler-Golden, S.M. (2012). Quick atmospheric correction code: Algorithm description and recent upgrades. *Optical Engineering, 51*, 12

Bigdeli, B., Amini Amirkolaee, H., & Pahlavani, P. (2018). DTM extraction under forest canopy using LiDAR data and a modified invasive weed optimization algorithm. *Remote Sensing of Environment, 216*, 289-300

Blan, L., & Butler, R. (1999). Comparing effects of aggregation methods on statistical and spatial properties of simulated spatial data. *Photogrammetric Engineering & Remote Sensing, 65*, 73-84

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 2-16

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., & Tiede, D. (2014). Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing, 87*, 180-191

Bocedi, G., Pe'er, G., Heikkinen, R.K., Matsinos, Y., & Travis, J.M.J. (2012). Projecting species' range expansion dynamics: sources of systematic biases when scaling up patterns and processes. *Methods in Ecology and Evolution, 3*, 1008-1018

Bolton, D.K., Coops, N.C., & Wulder, M.A. (2015). Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne lidar data. *Remote Sensing of Environment, 163*, 48-60

Bolton, D.K., White, J.C., Wulder, M.A., Coops, N.C., Hermosilla, T., & Yuan, X. (2018). Updating stand-level forest inventories using airborne laser scanning and Landsat time series data. *International Journal of Applied Earth Observation and Geoinformation, 66*, 174-183

Bonnet, S., Gaulton, R., Lehaire, F., & Lejeune, P. (2015). Canopy gap mapping from airborne laser scanning: An assessment of the positional and geometrical accuracy. *Remote Sensing, 7*, 11267-11294

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T., & Solberg, S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sensing of Environment, 114*, 911-924

Breiman, L. (1996). Bagging Predictors. *Machine Learning, 24*, 123–140

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5-32

Brodersen, K.H., Ong, C.S., Stephan, K.E., & Buhmann, J.M. (2010). The balanced accuracy and its posterior distribution. In, *2010 International Conference on Pattern Recognition* (pp. 3121-3124). Istanbul Convention & Exhibition Centre (ICEC), Istanbul, Turkey: IEEE

Brokaw, N.V.L., & Lent, R.A. (1999). Vertical structure. In M.L. Hunter (Ed.), *Maintaining Biodiversity in Forest Ecosystems* (pp. 373-399). Cambridge: Cambridge University Press

Brovkina, O., Hanuš, J., Zemek, F., Mochalov, V., Grigorieva, O., & Pikla, M. (2016). Evaluating the potential of satellite hyperspectral Resurs-P data for forest species classification. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B1*, 443-448

Brown, M.J., Elliott, H.J., & Hickey, J.E. (2001). An overview of the Warra Long-Term Ecological Research Site. *Tasforests, 13*, 1-8

Burai, P., Deák, B., Valkó, O., & Tomor, T. (2015). Classification of herbaceous vegetation using airborne hyperspectral imagery. *Remote Sensing, 7*, 2046-2066

Bureau of Meteorology (2017). Climate statistics for Australian locations (2004-2017). In. Australian Government, Australia

Bustamante, M.M., Roitman, I., Aide, T.M., Alencar, A., Anderson, L., Aragao, L., Asner, G.P., Barlow, J., Berenguer, E., Chambers, J., Costa, M.H., Fanin, T., Ferreira, L.G., Ferreira, J.N., Keller, M., Magnusson, W.E., Morales, L., Morton, D., Ometto, J.P., Palace, M., Peres, C., Silverio, D., Trumbore, S., & Vieira, I.C. (2015). Towards an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Global Change Biology, 22*, 92-109

Cadima, J., Cerdeira, J.O., & Minhoto, M. (2004). Computational aspects of algorithms for variable selection in the context of principal components. *Computational Statistics & Data Analysis, 47*, 225-236

Campbell, M.J., Dennison, P.E., Hudak, A.T., Parham, L.M., & Butler, B.W. (2018). Quantifying understory vegetation density using small-footprint airborne LiDAR. *Remote Sensing of Environment, 215*, 330-342

Camprodon, J., & Brotons, L. (2006). Effects of undergrowth clearing on the bird communities of the Northwestern Mediterranean Coppice Holm oak forests. *Forest Ecology and Management, 221*, 72-82

Cao, J., Leng, W., Liu, K., Liu, L., He, Z., & Zhu, Y. (2018). Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. *Remote Sensing, 10*, 89

Carlson, K.M., Asner, G.P., Hughes, R.F., Ostertag, R., & Martin, R.E. (2007). Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems, 10*, 536-549

Carmel, Y., Dean, D., & Flather, C. (2001). Combining location and classification error sources for estimating mult-temporal database accuracy. *Photogrammetric Engineering and Remote Sensing, 67*, 865-872

Casalí, J., Chahor, Y., Giménez, R., & Campo-Bescós, M. (2016). A critical discussion on the applicability of Compound Topographic Index (CTI) for predicting ephemeral gully erosion. *Geophysical Research Abstracts, 18*

Ceballos, A., Hernández, J., Corvalán, P., & Galleguillos, M. (2015). Comparison of airborne LiDAR and satellite hyperspectral remote sensing to estimate vascular plant richness in deciduous mediterranean forests of central Chile. *Remote Sensing, 7*, 2692-2714

Chan, J.C.-W., & Paelinckx, D. (2008). Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment, 112*, 2999-3011

Chang, K.T., Lin, C., Lin, Y.C., & Liu, J.K. (2016). Accuracy Assessment of Crown Delineation Methods for the Individual Trees Using Lidar Data. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B8*, 585-588

Chaube, N.R., Lele, N., Misra, A., Murthy, T., Manna, S., Hazra, S., Panda, M., & Samal, D.R. (2019). Mangrove species discrimination and health assessment using AVIRIS-NG hyperspectral data. *Current Science, 116*, 1136-1142

Chave, J. (2013). The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecol Lett, 16*, 4-16

Chen, D., Huang, J., & Jackson, T.J. (2005). Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing of Environment, 98*, 225-236

Chen, G., & Hay, G.J. (2011). An airborne LiDAR sampling strategy to model forest canopy height from Quickbird imagery and GEOBIA. *Remote Sensing of Environment, 115*, 1532-1542

Chen, G., Hay, G.J., & St-Onge, B. (2012). A GEOBIA framework to estimate forest parameters from lidar transects, Quickbird imagery and machine learning: A case study in Quebec, Canada. *International Journal of Applied Earth Observation and Geoinformation, 15*, 28-37

Chen, J., Gu, S., Shen, M., Tang, Y., & Matsushita, B. (2009). Estimating aboveground biomass of grassland having a high canopy cover: an exploratory analysis of in situ hyperspectral data. *International Journal of Remote Sensing, 30*, 6497-6517

Chen, J.M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing, 22*, 229-242

Cho, M.A., Mathieu, R., Asner, G.P., Naidoo, L., van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I.P.J., & Erasmus, B. (2012). Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment, 125*, 214-226

Cipriotti, P.A., Wiegand, T., Pütz, S., Bartoloni, N.J., & Paruelo, J.M. (2016). Nonparametric upscaling of stochastic simulation models using transition matrices. *Methods in Ecology and Evolution, 7*, 313-322

Clark, M., Roberts, D., & Clark, D. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment, 96*, 375-398

Clark, M.L., & Roberts, D.A. (2012). Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sensing, 4*, 1820-1855

Clawges, R., Vierling, K., Vierling, L., & Rowell, E. (2008). The use of airborne LiDAR to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sensing of Environment, 112*, 2064-2073

Clevers, J.G.P.W. (2014). Beyond NDVI: Extraction of biophysical variables from remote sensing imagery. In I. Manakos, & M. Braun (Eds.), *Land Use and Land Cover Mapping in Europe: Practices & Trends* (pp. 363-381). Dordrecht: Springer Netherlands

Clinton, N., Holt, A., Scarborough, J., Yan, L.I., & Gong, P. (2010). Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering and Remote Sensing, 76*, 289-299

Cody, M.L. (1981). Habitat selection in birds: The roles of vegetation structure, competitors, and productivity. *BioScience, 31*, 107-113

Cohen, W.B., & Goward, S. (2004). Landsat's role in ecological applications of remote sensing. *BioScience, 54*, 535-545

Cohen, W.B., & Spies, T.A. (1992). Estimating structural attributes of Douglas-fir/western hemlock forest stands from landsat and SPOT imagery. *Remote Sensing of Environment, 41*, 1-17

Colgan, M., Baldeck, C., Féret, J.-B., & Asner, G. (2012). Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. *Remote Sensing, 4*, 3462-3480

Colin, B., Schmidt, M., Clifford, S., Woodley, A., & Mengersen, K. (2018). Influence of spatial aggregation on prediction accuracy of green vegetation using boosted regression trees. *Remote Sensing, 10*, 1260

Coll, L., González-Olabarria, J.R., Mola-Yudego, B., Pukkala, T., & Messier, C. (2011). Predicting understory maximum shrubs cover using altitude and overstory basal area in different Mediterranean forests. *European Journal of Forest Research, 130*, 55-65

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development, 8*, 1991-2007

Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., & Trofymow, J.A. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees, 21*, 295

Corona, P., Chirici, G., & Marchetti, M. (2002). Forest ecosystem inventory and monitoring as a framework for terrestrial natural renewable resource survey programmes AU - Corona, Piermaria. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology, 136*, 69-82

Côté, J.-F., Widlowski, J.-L., Fournier, R.A., & Verstraete, M.M. (2009). The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. *Remote Sensing of Environment, 113*, 1067-1081

Crase, B., Vesk, P.A., Liedloff, A., & Wintle, B.A. (2015). Modelling both dominance and species distribution provides a more complete picture of changes to mangrove ecosystems under climate change. *Global Change Biology, 21*, 3005-3020

Criminisi, A., Shotton, J., & Konukoglu, E. (2011). Decision forests for classication, regression, density estimation, manifold learning and semi-supervised learning. In, *Microsoft Research technical report TR-2011-114* (p. 151). Cambridge, CB3 0FB, UK: Microsoft Research Ltd, 7 J J Thomson Ave, Cambridge, CB3 0FB, UK

Crippen, R.E. (1990). Calculating the vegetation index faster. *Remote Sensing of Environment, 34*, 71-73

Culbert, P., Radeloff, V., Flather, C., Kellndorfer, J., Rittenhouse, C., & Pidgeon, A. (2013). The influence of vertical and horizontal habitat structure on nationwide patterns of avian biodiversity. *The Auk, 130*, 656-665

Cutler, A., Cutler, D.R., & Stevens, J.R. (2012). Random Forests. In C. Zhang, & Y. Ma (Eds.), *Ensemble Machine Learning: Methods and Applications* (pp. 157-175). Boston, MA: Springer US

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., & Gibson, J.C. (2007). Random forests for classification in ecology. *Ecology, 88*, 2783-2792

Dalponte, M., Bruzzone, L., & Gianelle, D. (2008). Fusion of hyperspectral and LiDAR remote sensing data for classification of complex forest areas. *IEEE Transactions on Geoscience and Remote Sensing, 46*, 1416-1427

Dalponte, M., Bruzzone, L., & Gianelle, D. (2012). Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment, 123*, 258-270

Dalponte, M., Ørka, H.O., Ene, L.T., Gobakken, T., & Næsset, E. (2014). Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sensing of Environment, 140*, 306-317

Dalponte, M., Reyes, F., Kandare, K., & Gianelle, D. (2015). Delineation of individual tree crowns from ALS and hyperspectral data: A comparison among four methods. *European Journal of Remote Sensing, 48*, 365-382

Darwish, A., Leukert, K., & Reinhardt, W. (2003). Image segmentation for the purpose of object-based classification. In, *IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)* (pp. 2039-2041). Toulouse, France: IEEE

Dash, J.P., Watt, M.S., Bhandari, S., & Watt, P. (2016). Characterising forest structure using combinations of airborne laser scanning data, RapidEye satellite imagery and environmental variables. *Forestry, 89*, 159-169

Dean, C., Wardell-Johnson, G.W., & Kirkpatrick, J.B. (2012). Are there any circumstances in which logging primary wet-eucalypt forest will not add to the global carbon burden? *Agricultural and Forest Meteorology, 161*, 156-169

Debes, C., Merentitis, A., Heremans, R., Hahn, J., Frangiadakis, N., van Kasteren, T., Liao, W., Bellens, R., Pizurica, A., Gautama, S., Philips, W., Prasad, S., Du, Q., & Pacifici, F. (2014). Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 2405-2418

Degerickx, J., Roberts, D.A., McFadden, J.P., Hermy, M., & Somers, B. (2018). Urban tree health assessment using airborne hyperspectral and LiDAR imagery. *International Journal of Applied Earth Observation and Geoinformation, 73*, 26-38

Dehzangi, A., Heffernan, R., Sharma, A., Lyons, J., Paliwal, K., & Sattar, A. (2015). Gram-positive and Gram-negative protein subcellular localization by incorporating evolutionary-based descriptors into Chous general PseAAC. *Journal of Theoretical Biology, 364*, 284-294

Deng, Y., Wilson, J.P., & Bauer, B.O. (2007). DEM resolution dependencies of terrain attributes across a landscape. *International Journal of Geographical Information Science, 21*, 187-213

Desmet, P.J.J., & Govers, G. (1996). A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *Journal of Soil and Water Conservation, 51*, 427-433

Devriendt, F., Coilliea, F.V., Wulf, R.D., Bertels, L., Kempeneers, P., Vandekerkhove, K., & Morsdorf, F. (2012). Classification of unmanaged forest reserves in Flanders (Belgium) at the tree crown level using airborne hyperspectral and LiDAR data. In, *Proceedings of the 4th GEOBIA, May 7-9, 2012*. Rio de Janeiro, Brazil

Di Lallo, G., Mundhenk, P., Zamora López, S., Marchetti, M., & Köhl, M. (2017). REDD+: Quick assessment of deforestation risk based on available data. *Forests, 8*, 29

Dian, Y., Pang, Y., Dong, Y., & Li, Z. (2016). Urban tree species mapping using airborne LiDAR and hyperspectral data. *Journal of the Indian Society of Remote Sensing, 44*, 595-603

Dobrowski, S.Z., Safford, H.D., Cheng, Y.B., & Ustin, S.L. (2008). Mapping mountain vegetation using species distribution modeling, image-based texture analysis, and object-based classification. *Applied Vegetation Science, 11*, 499-508

Dorren, L.K.A., Maier, B., & Seijmonsbergen, A.C. (2003). Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *Forest Ecology and Management, 183*, 31-46

Dube, T., & Mutanga, O. (2015). Investigating the robustness of the new Landsat-8 Operational Land Imager derived texture metrics in estimating plantation forest aboveground biomass in resource constrained areas. *ISPRS Journal of Photogrammetry and Remote Sensing, 108*, 12-32

Dube, T., & Mutanga, O. (2016). The impact of integrating WorldView-2 sensor and environmental variables in estimating plantation forest species aboveground biomass and carbon stocks in uMgeni Catchment, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing, 119*, 415-425

Duro, D.C., Franklin, S.E., & Dubé, M.G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment, 118*, 259-272

Eckert, S. (2012). Improved forest biomass and carbon estimations using texture measures from WorldView-2 satellite data. *Remote Sensing, 4*, 810

eCognition (2014). Trimble eCognition® Developer 9.0 (User Guide). In (p. 262). Munich, Germany

Ehle, D.S., & Baker, W.L. (2003). Disturbance and stand dynamics in ponderosa pine forests in rocky mountain national park, USA. *Ecological Monographs, 73*, 543-566

Ehlers, S., Saarela, S., Lindgren, N., Lindberg, E., Nyström, M., Persson, H.J., Olsson, H., & Ståhl, G. (2018). Assessing Error Correlations in Remote Sensing-Based Estimates of Forest Attributes for Improved Composite Estimation. *Remote Sensing, 10*, 667

Elliott, H.J., Felton, K.C., Jarman, S.J., & Stone, M.G. (2008). A history of innovation: Eighty-five years of research and development at Forestry Tasmania. In (p. 402). Hobart, Tasmania: Forestry Tasmania

Erdody, T.L., & Moskal, L.M. (2010). Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment, 114*, 725-737

Eskelson, B.N.I., Madsen, L., Hagar, J.C., & Temesgen, H. (2011). Estimating riparian understory vegetation cover with beta regression and copula models. *Forest Science, 57*, 212-221

Falkowski, M.J., Evans, J.S., Martinuzzi, S., Gessler, P.E., & Hudak, A.T. (2009). Characterizing forest succession with LiDAR data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment, 113*, 946-956

Fassnacht, F.E., Latifi, H., Ghosh, A., Joshi, P.K., & Koch, B. (2014a). Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. *Remote Sensing of Environment, 140*, 533-548

Fassnacht, F.E., Neumann, C., Forster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., & Koch, B. (2014b). Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central european test sites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 2547-2561

Fekety, P.A., Falkowski, M.J., & Hudak, A.T. (2014). Temporal transferability of LiDAR-based imputation of forest inventory attributes. *Canadian Journal of Forest Research, 45*, 422-435

Féret, J.B., & Asner, G.P. (2012). Semi-supervised methods to identify individual crowns of lowland tropical canopy species using imaging spectroscopy and LiDAR. *Remote Sensing, 4*, 2457-2476

Ferreira, M.P., Zortea, M., Zanotta, D.C., Shimabukuro, Y.E., & de Souza Filho, C.R. (2016). Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sensing of Environment, 179*, 66-78

Ferris, R., & Humphrey, J.W. (1999). A review of potential biodiversity indicators for application in British forests. *Forestry: An International Journal of Forest Research, 72*, 313–328

Ferro, J.C., & Warner, T. (2002). Scale and texture in digital image classification. *Photogrammetric Engineering and Remote Sensing, 68*, 51-63

Filippelli, S.K., Lefsky, M.A., & Rocca, M.E. (2019). Comparison and integration of LiDAR and photogrammetric point clouds for mapping pre-fire forest structure. *Remote Sensing of Environment, 224*, 154-166

Flood, N. (2014). Continuity of reflectance data between Landsat-7 ETM+ and Landsat-8 OLI, for both top-of-atmosphere and surface reflectance: A study in the Australian landscape. *Remote Sensing, 6*, 7952

Forest Practices Authority (2017). State of the forests Tasmania 2017, booklet. In (p. 52). Hobart, Tasamania, Australia: Forest Practices Authority

Forestry Tasmania (2009a). Lowland wet eucalypt forests, Native Forest Silviculture, Technical Bulletin No. 8. In. 79 Melville Street, HOBART 7000, Australia: Forestry Tasmania

Forestry Tasmania (2009b). A new silviculture for Tasmania's public forests: A review of the variable retention program. In (p. 107). Hobart, Australia: Forestry Tasmania

Franklin, J. (1998). Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science, 9*, 733-748

Frazer, G.W., Wulder, M.A., & Niemann, K.O. (2005). Simulation and quantification of the fine-scale spatial pattern and heterogeneity of forest canopy structure: A lacunarity-based method designed for analysis of continuous canopy heights. *Forest Ecology and Management, 214*, 65-90

Freeman, E.A., Moisen, G.G., Coulston, J.W., & Wilson, B.T. (2015). Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. *Canadian Journal of Forest Research, 46*, 323-339

Friedman, J.H. (2001). 1999 Reitz Lecture Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics, 29*, 1189–1232

Fu, B.J., Liu, S.L., Ma, K.M., & Zhu, Y.G. (2004). Relationships between soil characteristics, topography and plant diversity in a heterogeneous deciduous broad-leaved forest near Beijing, China. *Plant and Soil, 261*, 47-54

Gao, L., Zhao, B., Jia, X., Liao, W., & Zhang, B. (2017). Optimized kernel minimum noise fraction transformation for hyperspectral image classification. *Remote Sensing, 9*

Gao, Y., Lu, D., Li, G., Wang, G., Chen, Q., Liu, L., & Li, D. (2018). Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. *Remote Sensing, 10*, 627

Garabedian, J.E., McGaughey, R.J., Reutebuch, S.E., Parresol, B.R., Kilgo, J.C., Moorman, C.E., & Peterson, M.N. (2014). Quantitative analysis of woodpecker habitat using high-resolution airborne LiDAR estimates of forest structure and composition. *Remote Sensing of Environment, 145*, 68-80

Gasparri, N.I., Parmuchi, M.G., Bono, J., Karszenbaum, H., & Montenegro, C.L. (2010). Assessing multi-temporal Landsat 7 ETM+ images for estimating above-ground biomass in subtropical dry forests of Argentina. *Journal of Arid Environments, 74*, 1262-1270

Ge, S., Carruthers, R., Gong, P., & Herrera, A. (2006). Texture analysis for mapping *Tamarix parviflora* using aerial photographs along the Cache Creek, California. *Environ Monit Assess, 114*, 65-83

Gebreslasie, M.T., Ahmed, F.B., & van Aardt, J.A.N. (2010). Predicting forest structural attributes using ancillary data and ASTER satellite data. *International Journal of Applied Earth Observation and Geoinformation, 12*, S23-S26

Geerling, G.W., Labrador-Garcia, M., Clevers, J.G.P.W., Ragas, A.M.J., & Smits, A.J.M. (2007). Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. *International Journal of Remote Sensing, 28*, 4263-4284

Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2015). VSURF: An R package for variable selection using random forests. *The R Journal, 7*, 19-33

George, R., Padalia, H., & Kushwaha, S.P.S. (2014). Forest tree species discrimination in western Himalaya using EO-1 Hyperion. *International Journal of Applied Earth Observation and Geoinformation, 28*, 140-149

Gessler, P.E., Moore, I.D., McKenzie, N.J., & Ryan, P.J. (1995). Soil-landscape modelling and spatial prediction of soil attributes. *International journal of geographical information systems, 9*, 421-432

Ghamisi, P., Benediktsson, J.A., & Phinn, S. (2014). Fusion of hyperspectral and LiDAR data in classification of urban areas. In, *International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 181-184). Quebec City Convention Centre Quebec City, QC, Canada

Ghosh, A., Fassnacht, F.E., Joshi, P.K., & Kochb, B. (2014). A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation, 26*, 49-63

Gilbert, J.M. (1959). Forest succession in the Florentine Valley, Tasmania In, *Papers and Proceedings of the Royal Society of Tasamania* (pp. 129-152)

Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P., & Saatchi, S. (2008). Measuring and modelling biodiversity from space. *Progress in Physical Geography, 32*, 203-221

Gillin, C.P., Bailey, S.W., McGuire, K.J., & Prisley, S.P. (2015). Evaluation of LiDAR-derived DEMs through terrain analysis and field comparison. *Photogrammetric Engineering & Remote Sensing, 81*, 387-396

Gitelson, A.A., Kaufman, Y.J., & Merzlyak, M.N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment, 58*, 289-298

Gitelson, A.A., & Merzlyak, M.N. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research, 22*, 689-692

Glenn, E.P., Huete, A.R., Nagler, P.L., & Nelson, S.G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors, 8*, 2136

Goel, N.S., & Qin, W. (1994). Influences of canopy architecture on relationships between various vegetation indices and LAI and Fpar: A computer simulation. *Remote Sensing Reviews, 10*, 309-347

Goetz, S., Steinberg, D., Dubayah, R., & Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment, 108*, 254-263

Goetz, S.J., Steinberg, D., Betts, M.G., Holmes, R.T., Doran, P.J., Dubayah, R., & Hofton, M. (2010). LiDAR remote sensing variables predict breeding habitat of a Neotropical migrant bird. *Ecology, 91*, 1569-1576

Gomez-Chova, L., Calpe, J., Soria, E., Camps-Valls, G., Martin, J.D., & Moreno, J. (2003). Cart-based feature selection of hyperspectral lmages for crop cover classification. In, *Proceedings of International Conference on Image Processing* (pp. 589-592). Barcelona, Spain, Spain: IEEE

Gómez, C., White, J.C., & Wulder, M.A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 116*, 55-72

Goodenough, D., Li, J., Asner, G., Schaepman, M., Ustin, S., & Dyk, A. (2006). Combining hyperspectral remote sensing and physical modeling for applications in land ecosystems. In, *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*. Denver, CO, USA: IEEE

Goodwin, N., Turner, R., & Merton, R. (2005). Classifying Eucalyptus forests with high spatial and spectral resolution imagery: An investigation of individual species and vegetation communities. *Australian Journal of Botany, 53*, 337-345

Goodwin, N.R., Coops, N.C., & Culvenor, D.S. (2006). Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sensing of Environment, 103*, 140-152

Goodwin, N.R., Coops, N.C., & Culvenor, D.S. (2007). Development of a simulation model to predict LiDAR interception in forested environments. *Remote Sensing of Environment, 111*, 481-492

Gracia, M., Montané, F., Piqué, J., & Retana, J. (2007). Overstory structure and topographic gradients determining diversity and abundance of understory shrub species in temperate forests in central Pyrenees (NE Spain). *Forest Ecology and Management, 242*, 391-397

Green, A.A., Berman, M., Switzer, P., & Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing, 26*, 65-74

Greenwell, B.M. (2017). pdp: An R package for constructing partial dependence plots. *The R Journal (contributed research article), 9/1*, 421-436

Hagen, S.C., Heilman, P., Marsett, R., Torbick, N., Salas, W., van Ravensway, J., & Qi, J. (2012). Mapping total vegetation cover across western rangelands with moderate-resolution imaging spectroradiometer data. *Rangeland Ecology & Management, 65*, 456-467

Hamraz, H., Contreras, M.A., & Zhang, J. (2017). Vertical stratification of forest canopy for segmentation of understory trees within small-footprint airborne LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing, 130*, 385-392

Haralick, R.M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*, 610-621

Hastie, T., Tibshirani, R., & Friedman, J. (2008). Elements of statistical learning: Data mining, inference, and prediction (second edition). In (p. 745). Stanford, California: Springer

Haywood, A., & Stone, C. (2011). Using airborne laser scanning data to estimate structural attributes of natural eucalypt regrowth forests. *Australian Forestry, 74*, 4-12

Heinzel, J., & Koch, B. (2011). Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation, 13*, 152-160

Hernández-Stefanoni, J.L., Johnson, K.D., Cook, B.D., Dupuy, J.M., Birdsey, R., Peduzzi, A., & Tun-Dzul, F. (2015). Estimating species richness and biomass of tropical dry forests using LIDAR during leaf-on and leaf-off canopy conditions. *Applied Vegetation Science, 18*, 724-732

Hickey, J., Neyland, M., Rothe, A., & Bauhus, J. (2015). Is continuous-cover silviculture, as practised in Bavaria, suitable for use in wet eucalypt forests in Tasmania, Australia? *Australian Forestry, 78*, 29-44

Hickey, J.E. (1994a). A floristic comparison of vascular species in Tasmanian oldgrowth mixed forest with regeneration resulting from logging and wildfire. *Australian Journal of Botany, 42*, 383-404

Hickey, J.E. (1994b). A Floristic Comparison of Vascular Species in Tasmanian Oldgrowth Mixed Forest with Regeneration Resulting from Logging and Wildfire. *Australian Journal of Botany, 42*, 383-404

Hickey, J.E., Neyland, M.G., & Bassett, O.D. (2001). Rationale and design for the Warra silvicultural systems trial in wet *Eucalyptus obliqua* forests in Tasmania. *Tasforests, 13*, 155-182

Hickey, J.E., Neyland, M.G., Grove, S., & Edwards, L.G. (2006). From little things big things grow: The Warra Silvicultural Systems Trial in Tasmanian wet *Eucalyptus obliqua* forest. *Allgemeine Forst und Jagdzeitung, 177*

Hickey, J.E., Su, W., Rowe, P., Brown, M.J., & Edwards, L. (1999). Fire history of the tall wet eucalypt forests of the Warra ecological research site, Tamania. *Australian Forestry, 62*, 66-71

Hickey, J.E., & Wilkinson, G.R. (1999). The development and current implementation of silvicultural pratices in native forests in Tasmania. *Australian Forestry, 62*, 245-254

Higgins, M.A., Asner, G.P., Martin, R.E., Knapp, D.E., Anderson, C., Kennedy-Bowdoin, T., Saenz, R., Aguilar, A., & Joseph Wright, S. (2014). Linking imaging spectroscopy and LiDAR with floristic composition and forest structure in Panama. *Remote Sensing of Environment, 154*, 358-367

Hill, R., & Hinsley, S. (2015). Airborne LiDAR for woodland habitat quality monitoring: Exploring the significance of LiDAR data characteristics when modelling organism-habitat relationships. *Remote Sensing, 7*, 3446-3466

Hill, R.A., & Broughton, R.K. (2009). Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*, 223-233

Hindrum, L., Hovenden, M.J., Neyland, M.G., & Baker, S.C. (2012). The effects of mechanical disturbance and burn intensity on the floristic composition of two-year old aggregated retention coupes in Tasmanian wet eucalypt forests. *Forest Ecology and Management, 279*, 55-65

Hinsley, S., Hill, R., Bellamy, P., & Balzter, H. (2006). The application of LiDAR in woodland bird ecology: Climate, canopy structure, and habitat quality. *Photogrammetric Engineering & Remote Sensing,, 72*, 1399-1406

Hirata, Y., Tabuchi, R., Patanaponpaiboon, P., Poungparn, S., Yoneda, R., & Fujioka, Y. (2014). Estimation of aboveground biomass in mangrove forests using high-resolution satellite data. *Journal of Forest Research, 19*, 34-41

Holmgren, J., & Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment, 90*, 415-423

Holopainen, M., Vastaranta, M., Karjalainen, M., Karila, K., Kaasalainen, S., Honkavaara, E., & Hyyppä, J. (2015). Forest Inventory Attribute Estimation Using Airborne Laser Scanning, Aerial Stereo Imagery, Radargrammetry and Interferometry–Finnish Experiences of the 3d Techniques. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3/W4*, 63-69

Hong, S.h., Hendrickx, J.M.H., & Borchers, B. (2009). Up-scaling of SEBAL derived evapotranspiration maps from Landsat (30 m) to MODIS (250 m) scale. *Journal of Hydrology, 370*, 122-138

Hu, B., Li, J., Jing, L., & Judah, A. (2014). Improving the efficiency and accuracy of individual tree crown delineation from high-density LiDAR data. *International Journal of Applied Earth Observation and Geoinformation, 26*, 145-155

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E., & Morgan, P. (2006). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing, 32*, 126-138

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., & Falkowski, M.J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment, 112*, 2232-2245

Hyde, P., Dubayah, R., Peterson, B., Blair, J.B., Hofton, M., Hunsaker, C., Knox, R., & Walker, W. (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment, 96*, 427-437

Imdadullah, M., Aslam, M., & Altaf, S. (2016). *mctest*: An R package for detection of collinearity among regressors. *The R Journal, 8*

Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree species classification with random forest using very high spatial resolution 8-Band WorldView-2 satellite data. *Remote Sensing, 4*, 2661-2693

Iqbal, I.A., Musk, R.A., Osborn, J., Stone, C., & Lucieer, A. (2019). A comparison of area-based forest attributes derived from airborne laser scanner, small-format and medium-format digital aerial photography. *International Journal of Applied Earth Observation and Geoinformation, 76*, 231-241

Jackson, R.D., & Huete, A.R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine, 11*, 185-200

Jackson, W.D. (1968). Fire, air, water and earth—an elemental ecology of Tasmania. In, *Proceedings of the Ecological Society of Australia* (pp. 9-16)

Jacoby, B.S., Peterson, E.W., & Dogwiler, T. (2011). Identifying the Stream Erosion Potential of Cave Levels in Carter Cave State Resort Park, Kentucky, USA. *Journal of Geographic Information System, 3*, 323-333

Jakubowski, M.K., Guo, Q., & Kelly, M. (2013). Tradeoffs between lidar pulse density and forest measurement accuracy. *Remote Sensing of Environment, 130*, 245-253

Jaskierniak, D., Lane, P.N.J., Robinson, A., & Lucieer, A. (2011). Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. *Remote Sensing of Environment, 115*, 573-585

Jayathunga, S., Owari, T., & Tsuyuki, S. (2018). Evaluating the performance of photogrammetric products using fixed-wing UAV imagery over a mixed conifer–broadleaf forest: Comparison with airborne laser scanning. *Remote Sensing, 10*, 187

Jenkins, R.B., & Coops, N.C. (2011). Landscape controls on structural variation in eucalypt vegetation communities: Woronora Plateau, Australia. *Australian Geographer, 42*, 1-17

Ji, L., Zhang, L., Wylie, B., & Rover, J. (2011). On the terminology of the spectral vegetation index (NIR − SWIR)/(NIR + SWIR). *International Journal of Remote Sensing, 32*, 6901-6909

Johansen, K., & Phinn, S. (2013). Remote sensing of biophysical parameters: Linking field, airborne and continental scale data. In: Presentation for Terrestrial Ecosystem Research Network, Australia.

Johansen, K., Phinn, S., & Witte, C. (2010). Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs. *Remote Sensing of Environment, 114*, 2679-2691

Jones, T.G., Coops, N.C., & Sharma, T. (2010). Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment, 114*, 2841-2852

Kamal, M., Phinn, S., & Johansen, K. (2014). Characterizing the spatial structure of mangrove features for optimizing image-based mangrove mapping. *Remote Sensing, 6*, 984

Kamal, M., Phinn, S., & Johansen, K. (2015). Object-based approach for multi-scale mangrove composition mapping using multi-resolution image datasets. *Remote Sensing, 7*, 4753-4783

Kasel, S., Bennett, L.T., Aponte, C., Fedrigo, M., & Nitschke, C.R. (2017). Environmental heterogeneity promotes floristic turnover in temperate forests of south-eastern Australia more than dispersal limitation and disturbance. *Landscape Ecology, 32*, 1613-1629

Kayitakire, F., Hamel, C., & Defourny, P. (2006). Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment, 102*, 390-401

Ke, Y., Quackenbush, L.J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LiDAR data for object-based forest species classification. *Remote Sensing of Environment, 114*, 1141-1154

Kemppinen, J., Niittynen, P., Riihimäki, H., & Luoto, M. (2018). Modelling soil moisture in a high-latitude landscape using LiDAR and soil data. *Earth Surface Processes and Landforms, 43*, 1019-1031

Key, C.H., & Benson, N.C. (2006). Landscape assessment: Ground measure of severity, the composite burn index; and remote sensing of severity, the normalized burn ratio. In R.E.K. D.C. Lutes, J.F. Caratti, C.H. Key, N.C. Benson, S. Sutherland, L.J. Gangi (Ed.),

FIREMON: Fire Effects Monitoring and Inventory System (p. 56). Ogden, UT: USDA Forest Service, Rocky Mountain Research Station, Ogden, UT

Khodadadzadeh, M., Prasad, S., & Plaza, A. (2015). Fusion of hyperspectral and LiDAR remote sensing data using multiple feature learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*, 2971-2983

Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., & Hussin, Y.A. (2014). Generating pit-free canopy height models from airborne lidar. *Photogrammetric Engineering and Remote Sensing, 80*, 863-872

Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS, 8*, 83-111

King, D.A., & Clark, D.A. (2011). Allometry of emergent tree species from saplings to above-canopy adults in a Costa Rican rain forest. *Journal of Tropical Ecology, 27*, 573-579

Kinnell, P. (2005). Alternative approaches for determining the USLE-M slope length factor for grid cells. *Soil Science Society of America Journal, 69*, 674-680

Kirkpatrick, J.B., Green, K., Bridle, K.L., & Venn, S.E. (2014). Patterns of variation in Australian alpine soils and their relationships to parent material, vegetation formation, climate and topography. *CATENA, 121*, 186-194

Kiss, R. (2004). Determination of drainage network in digital elevation models, utilities and limitations. *Journal of Hungarian Geomathematics, 2*, 16-29

Koch, A., Munks, S., & Driscoll, D. (2008). The use of hollow-bearing trees by vertebrate fauna in wet and dry Eucalyptus obliqua forest, Tasmania. *Wildlife Research, 35*, 727–746

Koetz, B., Morsdorf, F., van der Linden, S., Curt, T., & Allgöwer, B. (2008). Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management, 256*, 263-271

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C., Zhang, J., Dech, S., & Wikelski, M. (2014). Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks. *International Journal of Remote Sensing, 35*, 6599-6647

Kuester, M. (2016). Radiometric use of WorldView-3 imagery-Technical Note. In (pp. 1-12). 1601 Dry Creek Drive, Suite 260, Longmont CO 80503, USA: DigitalGlobe

Kuhn, M. (2017). Classification and Regression Training (Package '*caret*'). In (p. 205)

Kunin, W.E., Harte, J., He, F., Hui, C., Jobe, R.T., Ostling, A., Polce, C., Šizling, A., Smith, A.B., Smith, K., Smart, S.M., Storch, D., Tjørve, E., Ugland, K.-I., Ulrich, W., & Varma, V. (2018). Upscaling biodiversity: estimating the species–area relationship from small samples. *Ecological Monographs, 88*, 170-187

Kuyah, S., Muthuri, C., Jamnadass, R., Mwangi, P., Neufeldt, H., & Dietz, J. (2012). Crown area allometries for estimation of aboveground tree biomass in agricultural landscapes of western Kenya. *Agroforestry Systems, 86*, 267-277

La, P.H., Eo, Y.D., Nguyen, Q.M., & Kim, S.W. (2012). Individual tree crown estimation using hyperspectral image and LiDAR data. In, *7th International Conference on Computing and Convergence Technology (ICCCT)* (pp. 1413-1416): IEEE, Seoul

Laffan, M.D. (2001). Geology and soils of the Warra LTER Site: a preliminary description. *Tasforests, 13*, 23-30

Latifi, H. (2012). Characterizing forest structure by means of remote sensing: A review. In B. Escalante-Ramirez (Ed.), *Remote Sensing - Advanced Techniques and Platforms* (pp. 4–28). Zagreb, Croatia: Intech Open Access Publisher

Latifi, H., Fassnacht, F., & Koch, B. (2012a). Forest structure modeling with combined airborne hyperspectral and LiDAR data. *Remote Sensing of Environment, 121*, 10-25

Latifi, H., Heurich, M., Hartig, F., Müller, J., Krzystek, P., Jehl, H., & Dech, S. (2016). Estimating over- and understorey canopy density of temperate mixed stands by airborne LiDAR data. *Forestry, 89*, 69-81

Latifi, H., Hill, S., Schumann, B., Heurich, M., & Dech, S. (2017). Multi-model estimation of understorey shrub, herb and moss cover in temperate forest stands by laser scanner data. *Forestry: An International Journal of Forest Research, 90*, 496-514

Latifi, H., Nothdurft, A., Straub, C., & Koch, B. (2012b). Modelling stratified forest attributes using optical/LiDAR features in a central European landscape. *International Journal of Digital Earth, 5*, 106-132

Lawley, V., Lewis, M., Clarke, K., & Ostendorf, B. (2016). Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. *Ecological Indicators, 60*, 1273-1283

Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., & Shreenan, R. (2003). Combined high density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal of Remote Sensing, 29*, 633–649

Lee, A.C. (2008). Utilising airborne scanning laser (LiDAR) to improve the assessment of Australian native forest structure. In (p. 460). Canberra, Australia: PhD Thesis. Australian National University

Lee, A.C., & Lucas, R.M. (2007). A LiDAR-derived canopy density model for tree stem and crown mapping in Australian forests. *Remote Sensing of Environment, 111*, 493-518

Lefsky, M.A., Cohen, W.B., Park, G.G., & Harding, D.J. (2002). LiDAR remote sensing for ecosystem studies. *BioScience, 52*, 19-30

Leiterer, R., Furrer, R., Schaepman, M.E., & Morsdorf, F. (2015a). Forest canopy-structure characterization: A data-driven approach. *Forest Ecology and Management, 358*, 48-61

Leiterer, R., Torabzadeh, H., Furrer, R., Schaepman, M., & Morsdorf, F. (2015b). Towards automated characterization of canopy layering in mixed temperate forests using airborne laser scanning. *Forests, 6*, 4146-4167

Lesak, A.A., Radeloff, V.C., Hawbaker, T.J., Pidgeon, A.M., Gobakken, T., & Contrucci, K. (2011). Modeling forest songbird species richness using LiDAR-derived measures of forest structure. *Remote Sensing of Environment, 115*, 2823-2835

Leutner, B.F., Reineking, B., Müller, J., Bachmann, M., Beierkuhnlein, C., Dech, S., & Wegmann, M. (2012). Modelling forest α-diversity and floristic composition - on the added value of LiDAR plus hyperspectral remote sensing. *Remote Sensing, 4*, 2818-2845

Levick, S., Setterfield, S., Rossiter-Rachor, N., Hutley, L., McMaster, D., & Hacker, J. (2015). Monitoring the distribution and dynamics of an invasive grass in tropical savanna using airborne LiDAR. *Remote Sensing, 7*, 5117

Levin, S.A. (1992). The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture. *Ecology, 73*, 1943-1967

Li, W., Niu, Z., Gao, S., Huang, N., & Chen, H. (2014). Correlating the horizontal and vertical distribution of LiDAR point clouds with components of biomass in a *picea crassifolia* forest. *Forests, 5*, 1910-1930

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. In, *R News* (pp. 18-22)

Liaw, A., & Wiener, M. (2018). Breiman and Cutler's Random Forests for Classification and Regression. R Package 'randomForest', Version 4.6-14. In

Lindberg, E., Roberge, J.M., Johansson, T., & Hjältén, J. (2015). Can airborne laser scanning (ALS) and forest estimates derived from satellite images be used to predict abundance and species richness of birds and beetles in boreal forest? *Remote Sensing, 7*, 4233-4252

Lindenmayer, D.B., Franklin, J.F., Lõhmus, A., Baker, S.C., Bauhus, J., Beese, W., Brodie, A., Kiehl, B., Kouki, J., Pastur, G.M., Messier, C., Neyland, M., Palik, B., Sverdrup-Thygeson, A., Volney, J., Wayne, A., & Gustafsson, L. (2012). A major shift to the retention approach for forestry can help resolve some global forest sustainability issues. *Conservation Letters, 5*, 421-431

Lindenmayer, D.B., Mackey, B.G., Mullen, I.C., McCarthy, M.A., Gill, A.M., Cunningham, R.B., & Donnelly, C.F. (1999). Factors affecting stand structure in forests – are there climatic and topographic determinants? *Forest Ecology and Management, 123*, 55-63

Liu, L., Coops, N.C., Aven, N.W., & Pang, Y. (2017). Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sensing of Environment, 200*, 170-182

Liu, L., Pang, Y., Fan, W., Li, Z., & Li, M. (2011). Fusion of airborne hyperspectral and LiDAR data for tree species classification in the temperate forest of Northeast China. In, *19th International Conference on Geoinformatics* (p. 5). Main Hall of East China Normal University, Shanghai, China: IEEE

Liu, X., & Bo, Y. (2015). Object-based crop species classification based on the combination of airborne hyperspectral images and LiDAR data. *Remote Sensing, 7*, 922-950

Lopatin, J., Dolos, K., Hernández, H.J., Galleguillos, M., & Fassnacht, F.E. (2016). Comparing Generalized Linear Models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. *Remote Sensing of Environment, 173*, 200-210

Lopatin, J., Galleguillos, M., Fassnacht, F.E., Ceballos, A., & Hernández, J. (2015). Using a multistructural object-based LiDAR approach to estimate vascular plant richness in mediterranean forests with complex structure. *IEEE Geoscience and Remote Sensing Letters, 12*, 1008-1012

Louppe, G. (2014). Unders tanding random forests: from theory to practice. In (p. 225). Wallonia, Belgium: PhD Thesis. Faculty of Applied Sciences, Department of Electrical Engineering & Computer Science, University of Liège

Lu, D. (2006). The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing, 27*, 1297-1328

Lucas, R., Bunting, P., Paterson, M., & Chisholm, L. (2008). Classification of Australian forest communities using aerial photography, CASI and HyMap data. *Remote Sensing of Environment, 112*, 2088-2103

Luo, G., Chen, G., Tian, L., Qin, K., & Qian, S.-E. (2016a). Minimum Noise Fraction versus Principal Component Analysis as a preprocessing step for hyperspectral imagery denoising. *Canadian Journal of Remote Sensing, 42*, 106-116

Luo, S., Chen, J.M., Wang, C., Xi, X., Zeng, H., Peng, D., & Li, D. (2016b). Effects of LiDAR point density, sampling size and height threshold on estimation accuracy of crop biophysical parameters. *Optics Express, 24*, 11578-11593

Luo, S., Wang, C., Xi, X., Pan, F., Peng, D., Zou, J., Nie, S., & Qin, H. (2017). Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecological Indicators, 73*, 378-387

Luscombe, D.J., Anderson, K., Gatis, N., Wetherelt, A., Grand-Clement, E., & Brazier, R.E. (2015). What does airborne LiDAR really measure in upland ecosystems? *Ecohydrology, 8*, 584-594

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 130*, 277-293

MacArthur, R.H., & MacArthur, J.W. (1961). On Bird Species Diversity. *Ecology, 42*, 594-598

Magnussen, S., Næsset, E., & Gobakken, T. (2010). Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Remote Sensing of Environment, 114*, 700-712

Maltamo, M., Packalén, P., Yu, X., Eerikäinen, K., Hyyppä, J., & Pitkänen, J. (2005). Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *Forest Ecology and Management, 216*, 41-50

Marrs, J., & Ni-Meister, W. (2019). Machine learning techniques for tree species classification using co-registered LiDAR and hyperspectral data. *Remote Sensing, 11*, 819

Marsett, R.C., Qi, J., Heilman, P., Biedenbender, S.H., Carolyn Watson, M., Amer, S., Weltz, M., Goodrich, D., & Marsett, R. (2006). Remote sensing for grassland management in the arid southwest. *Rangeland Ecology & Management, 59*, 530-540

Martinuzzi, S., Vierling, L.A., Gould, W.A., Falkowski, M.J., Evans, J.S., Hudak, A.T., & Vierling, K.T. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment, 113*, 2533-2546

Maschler, J., Atzberger, C., & Immitzer, M. (2018). Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. *Remote Sensing, 10*, 1218

Masek, J.G., Hayes, D.J., Joseph Hughes, M., Healey, S.P., & Turner, D.P. (2015). The role of remote sensing in process-scaling studies of managed forest ecosystems. *Forest Ecology and Management, 355*, 109-123

Maselli, F., Chirici, G., Bottai, L., Corona, P., & Marchetti, M. (2005). Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images. *International Journal of Remote Sensing, 26*, 3781-3796

Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Bolton, D.K., Tompalski, P., & Bater, C.W. (2018). Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots. *Remote Sensing of Environment, 216*, 697-714

Matsuki, T., Yokoya, N., & Iwasaki, A. (2013). Fusion of hyperspectral and LiDAR data for tree species classification. In, *34th Asian Conference on Remote Sensing (ACRS), Bali, 20-24 October, 2013*. Bali

Matsuki, T., Yokoya, N., & Iwasaki, A. (2015). Hyperspectral tree species classification of Japanese complex mixed forest with the aid of LiDAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*, 2177-2187

McElhinny, C., Gibbons, P., Brack, C., & Bauhus, J. (2005). Forest and woodland stand structural complexity: Its definition and measurement. *Forest Ecology and Management, 218*, 1-24

McInerney, D.O., Suarez-Minguez, J., Valbuena, R., & Nieuwenhuis, M. (2010). Forest canopy height retrieval using LiDAR data, medium-resolution satellite imagery and kNN estimation in Aberfoyle, Scotland. *Forestry: An International Journal of Forest Research, 83*, 195-206

Melin, M., Matala, J., Mehtätalo, L., Pusenius, J., & Packalen, P. (2016). Ecological dimensions of airborne laser scanning — Analyzing the role of forest structure in moose habitat use within a year. *Remote Sensing of Environment, 173*, 238-247

Mellor, A., Haywood, A., Jones, S., & Wilkes, P. (2012). Forest classification using random forests with multisource remote sensing and ancillary GIS data. In, *16th Australasian Remote Sensing and Photogrammetry Conference* (pp. 40-44). Melbourne, Victoria, Australia: SSSI Remote Sensing and Photogrammetry Commission

Mellor, A., Haywood, A., Stone, C., & Jones, S. (2013). The performance of random forests in an operational setting for large area sclerophyll forest classification. *Remote Sensing, 5*, 2838

Merentitis, A., Debes, C., Heremans, R., & Frangiadakis, N. (2014). Automatic fusion and classification of hyperspectral and LiDAR data using random forests. In, *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International* (pp. 1245-1248). Quebec City, QC, Canada: IEEE

Michez, A., Piegay, H., Lisein, J., Claessens, H., & Lejeune, P. (2016). Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environ Monit Assess, 188*, 146

Mikita, T., Klimánek, M., & Miloš, C. (2013). Evaluation of airborne laser scanning data for tree parameters and terrain modelling in forest environment. *Acta Universitatis Agriculturae Et Silviculturae Mendelianae Brunensis, LXI* 1339-1347

Ming, D., Li, J., Wang, J., & Zhang, M. (2015). Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example. *ISPRS Journal of Photogrammetry and Remote Sensing, 106*, 28-41

Mirik, M., Ansley, R.J., Steddom, K., Jones, D., Rush, C., Michels, G., & Elliott, N. (2013). Remote distinction of a noxious weed (Musk Thistle: *CarduusNutans*) using airborne hyperspectral imagery and the Support Vector Machine classifier. *Remote Sensing, 5*, 612-630

Moffett, K.B., & Gorelick, S.M. (2013). Distinguishing wetland vegetation and channel features with object-based image segmentation. *International Journal of Remote Sensing, 34*, 1332-1354

Moore, I.D., Gessler, P.E., Nielsen, G.A., & Peterson, G.A. (1993). Soil Attribute Prediction Using Terrain Analysis. *Soil Science Society of America Journal, 57*

Moore, I.D., Grayson, R.B., & Ladson, A.R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes, 5*, 3-30

Moore, I.D., O'Loughlin, E.M., & Burch, G.J. (1988). A contour-based topographic model for hydrological and ecological applications. *Earth Surface Processes and Landforms, 13*, 305-320

Moroni, M., H. Kelley, T., & L. McLarin, M. (2010). Carbon in trees in Tasmanian state forest. *International Journal of Forestry Research, 2010*, 13

Morsdorf, F., Mårell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E., & Allgöwer, B. (2010). Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. *Remote Sensing of Environment, 114*, 1403-1415

Moskal, L., & Jakubauskas, M. (2013). Monitoring post disturbance forest regeneration with hierarchical object-based image analysis. *Forests, 4*, 808-829

Moudrý, V., Gdulová, K., Fogl, M., Klápště, P., Urban, R., Komárek, J., Moudrá, L., Štroner, M., Barták, V., & Solský, M. (2019). Comparison of leaf-off and leaf-on combined UAV imagery and airborne LiDAR for assessment of a post-mining site terrain and vegetation structure: Prospects for monitoring hazards and restoration success. *Applied Geography, 104*, 32-41

Mücher, C.A., Roupioz, L., Kramer, H., Bogers, M.M.B., Jongman, R.H.G., Lucas, R.M., Kosmidou, V.E., Petrou, Z., Manakos, I., Padoa-Schioppa, E., Adamo, M., & Blonda, P. (2015). Synergy of airborne LiDAR and Worldview-2 satellite imagery for landcover and habitat mapping: A BIO SOS-EODHaM case study for the Netherlands. *International Journal of Applied Earth Observation and Geoinformation, 37*, 48-55

Mui, A., He, Y., & Weng, Q. (2015). An object-based approach to delineate wetlands across landscapes of varied disturbance with high spatial resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 109*, 30-46

Müller, J., & Brandl, R. (2009). Assessing biodiversity by remote sensing in mountainous terrain: The potential of LiDAR to predict forest beetle assemblages. *Journal of Applied Ecology, 46*, 897-905

Müller, J., Stadler, J., & Brandl, R. (2010). Composition versus physiognomy of vegetation as predictors of bird assemblages: The role of lidar. *Remote Sensing of Environment, 114*, 490-495

Mura, M., McRoberts, R.E., Chirici, G., & Marchetti, M. (2015). Estimating and mapping forest structural diversity using airborne laser scanning data. *Remote Sensing of Environment, 170*, 133-142

Murakami, T., Kawai, T., Mochizuki, S., Imai, Y., Kamagata, N., & Akamatsu, Y. (2011). Tree species classification using hyperspectral image combined with digital canopy height model. In, *32nd Asian Conference on Remote Sensing 2011, ACRS 2011* (pp. 894-899)

Murphy, P.N.C., Ogilvie, J., Meng, F.-R., White, B., Bhatti, J.S., & Arp, P.A. (2011). Modelling and mapping topographic variations in forest soils at high resolution: A case study. *Ecological Modelling, 222*, 2314-2332

Musk, R. (2017). Prospects for aerial photogrammetry canopy surface models in hardwood plantation inventory. In C.S.a.J. Osborn (Ed.), *Deployment and integration of cost-effective high resolution remotely sensed data for the Australian forest industry* (pp. 82-90). Queen Street Melbourne VIC 3000, Australia: Forest & Wood Products Australia

Nagendra, H. (2001). Using remote sensing to assess biodiversity. *International Journal of Remote Sensing, 22*, 2377-2400

Naidoo, L., Cho, M.A., Mathieu, R., & Asner, G. (2012). Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS Journal of Photogrammetry and Remote Sensing, 69*, 167-179

Neyland, M.G. (2001). Vegetation of the Warra silvicultural systems trial. *Tasforests, 13*, 183-192

Neyland, M.G., Hickey, J., Beadle, C., Bauhus, J., Davidson, N., & Edwards, L. (2009). An examination of stocking and early growth in the Warra silvicultural systems trial confirms the importance of a burnt seedbed for vigorous regeneration in Eucalyptus obliqua forest. *Forest Ecology and Management, 258*, 481-494

Neyland, M.G., Hickey, J., & Read, S.M. (2012). A synthesis of outcomes from the Warra Silvicultural Systems Trial, Tasmania: safety, timber production, economics, biodiversity, silviculture and social acceptability. *Australian Forestry, 75*, 147-162

Nielsen, A.A. (2011). Kernel maximum autocorrelation factor and minimum noise fraction transformations. *IEEE Transactions on Image Processing, 20*, 612-624

Nijland, W., Nielsen, S.E., Coops, N.C., Wulder, M.A., & Stenhouse, G.B. (2014). Fine-spatial scale predictions of understory species using climate- and LiDAR-derived terrain and canopy metrics. *Journal of Applied Remote Sensing, 8*, 16

Odom, R., & Henry McNab, W. (2000). Using digital terrain modeling to predict ecological types in the Balsam mountains of Western North Carolina. In (p. 13). United States of America: Department of Agriculture, Forest Service, Southern Research Station

Onojeghuo, A.O., & Onojeghuo, A.R. (2017). Object-based habitat mapping using very high spatial resolution multispectral and hyperspectral imagery with LiDAR data. *International Journal of Applied Earth Observation and Geoinformation, 59*, 79-91

Oumar, Z., & Mutanga, O. (2014). Integrating environmental variables and WorldView-2 image data to improve the prediction and mapping of Thaumastocoris peregrinus (bronze bug) damage in plantation forests. *ISPRS Journal of Photogrammetry and Remote Sensing, 87*, 39-46

Ozdemir, I., & Karnieli, A. (2011). Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *International Journal of Applied Earth Observation and Geoinformation, 13*, 701-710

Packalén, P., & Maltamo, M. (2006). Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. *Forest Science, 52*, 611-622

Parent, J.R., Volin, J.C., & Civco, D.L. (2015). A fully-automated approach to land cover mapping with airborne LiDAR and high resolution multispectral imagery in a forested suburban landscape. *ISPRS Journal of Photogrammetry and Remote Sensing, 104*, 18-29

Pasher, J., & King, D.J. (2010). Development of a forest structural complexity index based on multispectral airborne remote sensing and topographic data. *Canadian Journal of Forest Research, 41*, 44-58

Patenaude, G., Milne, R., & Dawson, T.P. (2005). Synthesis of remote sensing approaches for forest carbon estimation: reporting to the Kyoto Protocol. *Environmental Science and Policy, 8*, 161-178

Pearse, G.D., Watt, M.S., Dash, J.P., Stone, C., & Caccamo, G. (2019). Comparison of models describing forest inventory attributes using standard and voxel-based lidar predictors across a range of pulse densities. *International Journal of Applied Earth Observation and Geoinformation, 78*, 341-351

Pebesma, E., Bivand, R., Rowlingson, B., Gomez-Rubio, V., Hijmans, R.J., Sumner, M., MacQueen, D., Lemon, J., O'Brien, J., & O'Rourke, J. (2018). *sp*: Classes and methods for spatial data, Version 1.3-1. In (p. 118)

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Hoft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P., Stuart, S.N., Turak, E., Walpole, M., & Wegmann, M. (2013). Essential biodiversity variables. *Science, 339*, 277-278

Pesonen, A., Maltamo, M., Eerikäinen, K., & Packalèn, P. (2008). Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *Forest Ecology and Management, 255*, 3288-3296

Pham, L.T.H., Brabyn, L., & Ashraf, S. (2016). Combining QuickBird, LiDAR, and GIS topography indices to identify a single native tree species in a complex landscape using an object-based classification approach. *International Journal of Applied Earth Observation and Geoinformation, 50*, 187-197

Phua, M.-H., Johari, S.A., Wong, O.C., Ioki, K., Mahali, M., Nilus, R., Coomes, D.A., Maycock, C.R., & Hashim, M. (2017). Synergistic use of Landsat 8 OLI image and airborne LiDAR data for above-ground biomass estimation in tropical lowland rainforests. *Forest Ecology and Management, 406*, 163-171

Popescu, S.C., & Wynne, R.H. (2004). Seeing the trees in the forest: Using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing, 70*, 589-604

Pouliot, D.A., King, D.J., Bell, F.W., & Pitt, D.G. (2002). Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment, 82*, 322-334

Prasad, A.M., Iverson, L.R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems, 9*, 181-199

Priyadarshini, K.N., Sivashankari, V., Shekhar, S., & Balasubramani, K. (2019). Assessment on the Potential of Multispectral and Hyperspectral Datasets for Land Use/Land Cover Classification. *Proceedings, 24*, 12

Puttonen, E., Suomalainen, J., Hakala, T., Räikkönen, E., Kaartinen, H., Kaasalainen, S., & Litkey, P. (2010). Tree species classification from fused active hyperspectral reflectance and LiDAR measurements. *Forest Ecology and Management, 260*, 1843-1852

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment, 48*, 119-126

Qian, W., Yang, Y., & Zou, H. (2016). Tweedie's Compound Poisson Model with grouped elastic net. *Journal of Computational and Graphical Statistics, 25*, 606-625

R Core Team (2017). R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing, Vienna, Austria

Radoux, J., Bogaert, P., Fasbender, D., & Defourny, P. (2011). Thematic accuracy assessment of geographic object-based image classification. *International Journal of Geographical Information Science, 25*, 895-911

Räsänen, A., Kuitunen, M., Tomppo, E., & Lensu, A. (2014). Coupling high-resolution satellite imagery with ALS-based canopy height model and digital elevation model in object-based boreal forest habitat type classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 94*, 169-182

Reese, H., Nyström, M., Nordkvist, K., & Olsson, H. (2014). Combining airborne laser scanning data and optical satellite data for classification of alpine vegetation. *International Journal of Applied Earth Observation and Geoinformation, 27*, 81-90

Ridolfi, L., Laio, F., & D'Odorico, P. (2008). Fertility island formation and evolution in dryland ecosystems. *Ecology and Society, 13*, 5

Riedler, B., Pernkopf, L., Strasser, T., Lang, S., & Smith, G. (2015). A composite indicator for assessing habitat quality of riparian forests derived from Earth observation data. *International Journal of Applied Earth Observation and Geoinformation, 37*, 114-123

Riley, S., Degloria, S., & Elliot, S.D. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *International Journal of Science, 5*, 23-27

Rocchini, D., Hernández-Stefanoni, J.L., & He, K.S. (2015). Advancing species diversity estimate by remotely sensed proxies: A conceptual review. *Ecological Informatics, 25*, 22-28

Rocha, A., Groen, T., Skidmore, A., Darvishzadeh, R., & Willemen, L. (2018). Machine learning using hyperspectral data inaccurately predicts plant traits under spatial dependency. *Remote Sensing, 10*, 1263

Roll, U., Geffen, E., & Yom-Tov, Y. (2015). Linking vertebrate species richness to tree canopy height on a global scale. *Global Ecology and Biogeography, 24*, 814-825

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55*, 95-107

Roth, K.L., Roberts, D.A., Dennison, P.E., Peterson, S.H., & Alonzo, M. (2015). The impact of spatial resolution on the classification of plant species and functional types within imaging spectrometer data. *Remote Sensing of Environment, 171*, 45-57

Roujean, J.-L., & Breon, F.-M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment, 51*, 375-384

Rouse, J., Haas, R., Schell, J., & Deering, D. (1973). Monitoring vegetation systems in the great plains with ERTS. In, *Third ERTS Symposium, NASA* (pp. 309-317). United States: NASA

Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., & Zhang, C. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment, 114*, 35-49

Ruiz, L.Á., Recio, J.A., Crespo-Peremarch, P., & Sapena, M. (2016). An object-based approach for mapping forest structural types based on low-density LiDAR and multispectral imagery. *Geocarto International*, 1-15

Rutten, G., Ensslin, A., Hemp, A., & Fischer, M. (2015). Vertical and horizontal vegetation structure across natural and modified habitat types at Mount Kilimanjaro. *PLoS ONE, 10*, e0138822

Scanlan, I., McElhinny, C., & Turner, P. (2010). A methodology for modelling canopy structure: An exploratory analysis in the tall wet eucalypt forests of Southern Tasmania. *Forests, 1*, 4

Schroeder, T.A., Cohen, W.B., & Yang, Z. (2007). Patterns of forest regrowth following clearcutting in western Oregon as determined from a Landsat time-series. *Forest Ecology and Management, 243*, 259-273

Scott, R.E., Neyland, M.G., & Hovenden, M.J. (2015). Variable-retention harvesting in Tasmania: regeneration success? *Australian Forestry, 78*, 232-242

Scott, R.E., Neyland, M.G., & McElwee, D.J. (2013). Early regeneration results following aggregated retention harvesting of wet eucalypt forests in Tasmania, Australia. *Forest Ecology and Management, 302*, 254-263

Scott, R.E., Neyland, M.G., McElwee, D.J., & Baker, S.C. (2012). Burning outcomes following aggregated retention harvesting in old-growth wet eucalypt forests. *Forest Ecology and Management, 276*, 165-173

Shang, X. (2013). Evaluating the capability of machine-learning algorithms and object-oriented classification techniques using hyperspectral remote sensing for the discrimination of Australian native forest species in south-eastern Australia. In, *School of Earth and Environment Sciences* (p. 229). Australia: PhD Thesis. University of Wollongong

Shang, X., & Chisholm, L.A. (2014). Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 2481-2489

Sharma, A.K., Malik, A., & Rohilla, R. (2016). A robust mean shift integrating color, GLCM based texture features and frame differencing. *International Journal of Scientific & Engineering Research, 7*, 1386-1398

Shen, X., & Cao, L. (2017). Tree-species classification in subtropical forests using airborne hyperspectral and LiDAR data. *Remote Sensing, 9*, 1180

Shi, Y., Skidmore, A.K., Wang, T., Holzwarth, S., Heiden, U., Pinnel, N., Zhu, X., & Heurich, M. (2018). Tree species classification using plant functional traits from LiDAR and hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation, 73*, 207-219

Silva, C., Hudak, A., Vierling, L., Klauberg, C., Garcia, M., Ferraz, A., Keller, M., Eitel, J., & Saatchi, S. (2017). Impacts of airborne LiDAR pulse density on estimating biomass stocks and changes in a selectively logged tropical forest. *Remote Sensing, 9*

Simonson, W.D., Allen, H.D., & Coomes, D.A. (2012). Use of an airborne LiDAR system to model plant species composition and diversity of mediterranean oak forests. *Conservation Biology, 26*, 840-850

Simonson, W.D., Allen, H.D., & Coomes, D.A. (2014a). Applications of airborne LiDAR for the assessment of animal species diversity. *Methods in Ecology and Evolution, 5*, 719-729

Simonson, W.D., Allen, H.D., & Coomes, D.A. (2014b). Overstorey and topographic effects on understories: Evidence for linkage from cork oak (Quercus suber) forests in southern Spain. *Forest Ecology and Management, 328*, 35-44

Singh, J.S., Roy, P.S., Murthy, M.S.R., & Jha, C.S. (2010). Application of landscape ecology and remote sensing for assessment, monitoring and conservation of biodiversity. *Journal of the Indian Society of Remote Sensing, 38*, 365-385

Singh, K.K., Davis, A.J., & Meentemeyer, R.K. (2015). Detecting understory plant invasion in urban forests using LiDAR. *International Journal of Applied Earth Observation and Geoinformation, 38*, 267-279

Skidmore, A.K., Pettorelli, N., Mücher, C.A., & Wegmann, M. (2015). Agree on biodiversity metrics to track from Space. *Nature, 523*, 403-405

Soh, L.-K., & Tsatsoulis, C. (1999). Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing, 37*, 780-795

Soille, P. (2004). Optimal removal of spurious pits in grid digital elevation models. *Water resources research, 40*

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., & Macomber, S.A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment, 75*, 230-244

Sripada, R.P., Heiniger, R.W., White, J.G., & Meijer, A.D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal, 98*, 968-977

Stone, M.G. (1998). Forest-type mapping by photointerpretation: A multi-purpose base for Tasmania's forest management. *Tasforests, 10*, 15-32

Stratoulias, D., Balzter, H., Zlinszky, A., & Tóth, V.R. (2018). A comparison of airborne hyperspectral-based classifications of emergent wetland vegetation at Lake Balaton, Hungary. *International Journal of Remote Sensing, 39*, 5689-5715

Suchar, V.A., & Crookston, N.L. (2010). Understory cover and biomass indices predictions for forest ecosystems of the Northwestern United States. *Ecological Indicators, 10*, 602-609

Sugumaran, R., & Voss, M. (2007). Object-oriented classification of LiDAR-fused hyperspectral imagery for tree species identification in an urban environment. In, *IEEE, Urban Remote Sensing Joint Event* (p. 6). Paris, France: IEEE

Sumnall, M., Peduzzi, A., Fox, T.R., Wynne, R.H., Thomas, V.A., & Cook, B. (2016). Assessing the transferability of statistical predictive models for leaf area index between two airborne discrete return LiDAR sensor designs within multiple intensely managed Loblolly pine forest locations in the south-eastern USA. *Remote Sensing of Environment, 176*, 308-319

Sverdrup-Thygeson, A., Ørka, H.O., Gobakken, T., & Næsset, E. (2016). Can airborne laser scanning assist in mapping and monitoring natural forests? *Forest Ecology and Management, 369*, 116-125

Swatantran, A., Dubayah, R., Goetz, S., Hofton, M., Betts, M.G., Sun, M., Simard, M., & Holmes, R. (2012). Mapping migratory bird prevalence using remote sensing data fusion. *PLoS ONE, 7*, 11

Tabor, J., McElhinny, C., Hickey, J., & Wood, J. (2007). Colonisation of clearfelled coupes by rainforest tree species from mature mixed forest edges, Tasmania, Australia. *Forest Ecology and Management, 240*, 13-23

Tagliabue, G., Panigada, C., Colombo, R., Fava, F., Cilia, C., Baret, F., Vreys, K., Meuleman, K., & Rossini, M. (2016). Forest species mapping using airborne hyperspectral APEX data. *Miscellanea Geographica, 20*, 28-33

Temesgen, H., Affleck, D., Poudel, K., Gray, A., & Sessions, J. (2015). A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. *Scandinavian Journal of Forest Research, 30*, 326-335

Tewari, S., Kulhavý, J., Rock, B.N., & Hadaš, P. (2003). Remote monitoring of forest response to changed soil moisture regime due to river regulation. *Journal of Forest Science, 49*, 429-438

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M.C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of Biogeography, 31*, 79-92

Thomas, I.A., Jordan, P., Mellander, P.E., Fenton, O., Shine, O., Ó hUallacháin, D., Creamer, R., McDonald, N.T., Dunlop, P., & Murphy, P.N.C. (2016). Improving the identification of hydrologically sensitive areas using LiDAR DEMs for the delineation and mitigation of critical source areas of diffuse pollution. *Science of The Total Environment, 556*, 276-290

Thomas, I.A., Jordan, P., Shine, O., Fenton, O., Mellander, P.E., Dunlop, P., & Murphy, P.N.C. (2017). Defining optimal DEM resolutions and point densities for modelling hydrologically sensitive areas in agricultural catchments dominated by

microtopography. *International Journal of Applied Earth Observation and Geoinformation, 54*, 38-52

Thomas, S.C., Halpern, C.B., Falk, D.A., Liguori, D.A., & Austin, K.A. (1999). Plant diversity in managed forests: Understory responses to thinning and fertilization. *Ecological Applications, 9*, 864-879

Thuillier, G., Hersé, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., Simon, P.C., & Mandel, H. (2003). The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC Spectrometer from the Atlas and Eureca Missions. *Solar Physics, 214*, 1-22

Tickle, P.K., Lee, A., Lucas, R.M., Austin, J., & Witte, C. (2006). Quantifying Australian forest floristics and structure using small footprint LiDAR and large scale aerial photography. *Forest Ecology and Management, 223*, 379-394

Tng, D.Y.P., Williamson, G.J., Jordan, G.J., & Bowman, D.M.J.S. (2012). Giant eucalypts – globally unique fire-adapted rain-forest trees? *New Phytologist, 196*, 1001-1014

Tochon, G., Féret, J.B., Valero, S., Martin, R.E., Knapp, D.E., Salembier, P., Chanussot, J., & Asner, G.P. (2015). On the use of binary partition trees for the tree crown segmentation of tropical rainforest hyperspectral images. *Remote Sensing of Environment, 159*, 318-331

Tompalski, P., Coops, N.C., White, J.C., Wulder, M.A., & Pickell, P.D. (2015). Estimating forest site productivity using airborne laser scanning data and Landsat Time Series. *Canadian Journal of Remote Sensing, 41*, 232-245

Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., & Katila, M. (2008). Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment, 112*, 1982-1999

Torabzadeh, H., Morsdorf, F., Leiterer, R., & Schaepman, M. (2014). Fusing imaging spectrometry and airborne laser scanning data for tree species discrimination. In, *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*. Quebec City, QC: IEEE

Torbick, N., Ledoux, L., Salas, W., & Zhao, M. (2016). Regional mapping of plantation extent using multisensor imagery. *Remote Sensing, 8*, 236

Travis, M.R., Elsner, G.H., Iverson, W.D., & Jonnson, C.G. (1975). VIEWIT: computation of seen areas, slope, and aspect for landuse planning. In (p. 67). Forest Service, U.S.Department of Agriculture, P. 0.BOX 245, Berkeley, California 94701: USDA Forest Service, General Technical Report, PSW-1111975, 70. Pacific Southwest Forest and Range Exp. Stn., Berkeley, California

Trier, Ø.D., Salberg, A.-B., Kermit, M., Rudjord, Ø., Gobakken, T., Næsset, E., & Aarsten, D. (2018). Tree species classification in Norway from airborne hyperspectral and airborne laser scanning data. *European Journal of Remote Sensing, 51*, 336-351

Tuanmu, M.-N., Viña, A., Bearer, S., Xu, W., Ouyang, Z., Zhang, H., & Liu, J. (2010). Mapping understory vegetation using phenological characteristics derived from remotely sensed data. *Remote Sensing of Environment, 114*, 1833-1844

Tucker, C., Vanpraet, C., Gaston, A., & Boerwinkel, E. (1984). Regional analysis from data from heterogeneous pixels - Remote sensing of total dry matter production in the Senegalese Sahel

Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*, 127-150

Turner, D., Lucieer, A., Malenovský, Z., King, D., & Robinson, S.A. (2018). Assessment of Antarctic moss health from multi-sensor UAS imagery with random forest modelling. *International Journal of Applied Earth Observation and Geoinformation, 68*, 168-179

Turner, P.A.M., Balmer, J., & Kirkpatrick, J.B. (2009). Stand-replacing wildfires?: The incidence of multi-cohort and single-cohort Eucalyptus regnans and E. obliqua forests in southern Tasmania. *Forest Ecology and Management, 258*, 366-375

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., & Woodcock, C. (2015). Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation, 182*, 173-176

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution, 18*, 306-314

Tyralis, H., & Papacharalampous, G. (2017). Variable selection in time series forecasting using random forests. *Algorithms, 10*, 114

USGS (2016). LANDSAT 8 (L8) Data Users Handbook, LSDS-1574, Version 2.0. In (p. 106). Sioux Falls, South Dakota, USA: Department of the Interior, U.S. Geological Survey

Vaglio Laurin, G., Chan, J.C.-W., Chen, Q., Lindsell, J.A., Coomes, D.A., Guerriero, L., Frate, F.D., Miglietta, F., & Valentini, R. (2014). Biodiversity mapping in a tropical West African forest with airborne hyperspectral data. *PLoS ONE, 9*, e97910

van Galen, L.G., Jordan, G.J., Musk, R.A., Beeton, N.J., Wardlaw, T.J., & Baker, S.C. (2018). Quantifying floristic and structural forest maturity: An attribute-based method for wet eucalypt forests. *Journal of Applied Ecology, 55*, 1668-1681

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G.P.W., & Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 108*, 273-290

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., & Clawges, R.M. (2008). LiDAR: shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment, 6*, 90-98

Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment, 115*, 3468-3478

Voss, M., & Sugumaran, R. (2008). The seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. *Sensors, 8*, 3020-3036

Wallace, L., Lucieer, A., Malenovský, Z., Turner, D., & Vopěnka, P. (2016). Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and Structure from Motion (SfM) point clouds. *Forests, 7*

Wallace, L., Lucieer, A., & Watson, C.S. (2014). Evaluating tree detection and segmentation routines on very high resolution UAV LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 7619-7628

Wallner, A., Elatawneh, A., Knoke, T., & Schneider, T. (2014). Estimation of forest structural information using RapidEye satellite data. *Forestry: An International Journal of Forest Research, 88*, 96-107

Wang, B., Zhang, G., & Duan, J. (2015). Relationship between topography and the distribution of understory vegetation in a Pinus massoniana forest in Southern China. *International Soil and Water Conservation Research, 3*, 291-304

Wang, H.-J., Prisley, S.P., Radtke, P.J., & Coulston, J. (2011). Errors in terrain-based model predictions caused by altered forest inventory plot locations in the southern appalachian mountains, USA. *Mathematical and Computational Forestry & Natural-Resource Sciences, 3*, 114–123

Wang, Z., Wu, J., Wang, Y., Kong, X., Bao, H., Ni, Y., Ma, L., & Jin, J. (2018). Crown-level tree species classification using integrated airborne hyperspectral and LiDAR remote sensing data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-3*, 2629-2634

Waske, B., Benediktsson, J.A., Árnason, K., & Sveinsson, J.R. (2009). Mapping of hyperspectral AVIRIS data using machine-learning algorithms. *Canadian Journal of Remote Sensing, 35*, S106-S116

Weisberg, P.J., Hadorn, C., & Bugmann, H. (2003). Predicting understorey vegetationcover from overstorey attributes in two temperate mountainforests. *Forstwissenschaftliches Centralblatt vereinigt mit Tharandter forstliches Jahrbuch, 122*, 273-286

Whiteside, T.G., Boggs, G.S., & Maier, S.W. (2011). Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation, 13*, 884-893

Wickham, H. (2017). *ggplot2: Elegant Graphics for Data Analysis. R package version 4.3.1.*: Springer-Verlag New York

Wickham, H., Hester, J., & Francois, R. (2017). readr: Read Tabular Data. R package version 1.1.1, https://cran.r-project.org/web/packages/readr/readr.pdf. In

Wilkes, P., Jones, S., Suarez, L., Mellor, A., Woodgate, W., Soto-Berelov, M., Haywood, A., & Skidmore, A. (2015). Mapping forest canopy height across large areas by upscaling ALS estimates with freely available satellite data. *Remote Sensing, 7*, 12563

Wilkes, P., Jones, S.D., Suarez, L., Haywood, A., Mellor, A., Woodgate, W., Soto-Berelov, M., Skidmore, A.K., & McMahon, S. (2016). Using discrete-return airborne laser scanning to quantify number of canopy strata across diverse forest types. *Methods in Ecology and Evolution, 7*, 700-712

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., & Grehan, A.J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy, 30*, 3-35

Wing, B.M., Ritchie, M.W., Boston, K., Cohen, W.B., Gitelman, A., & Olsen, M.J. (2012). Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment, 124*, 730-741

Winter, S., Chirici, G., McRoberts, R.E., Hauk, E., & Tomppo, E. (2008). Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. *Forestry: An International Journal of Forest Research, 81*, 33-44

Wood, E.M., Pidgeon, A.M., Radeloff, V.C., & Keuler, N.S. (2012). Image texture as a remotely sensed measure of vegetation structure. *Remote Sensing of Environment, 121*, 516-526

Wood, S., Stephens, H., & Bowman, D. (2015a). Ausplots forest monitoring network: Plot establishment report (Version 1.0). In (p. 90). Terrestrial Ecosystem Research Network, AusPlots Facility, Australia

Wood, S., Wardlaw, T., Pryde, E., & Baker, S. (2017). Landscape structure and mature forest biodiversity in wet eucalypt forests: A spatial analysis of timber production areas in South-Eastern Australia. *Forests, 8*, 89

Wood, S.W., & Bowman, D.M.J.S. (2012). Alternative stable states and the role of fire–vegetation–soil feedbacks in the temperate wilderness of southwest Tasmania. *Landscape Ecology, 27*, 13-28

Wood, S.W., Murphy, B.P., & Bowman, D.M.J.S. (2011). Firescape ecology: how topography determines the contrasting distribution of fire and rain forest in the south-

west of the Tasmanian Wilderness World Heritage Area. *Journal of Biogeography, 38*, 1807-1820

Wood, S.W., Prior, L.D., Stephens, H.C., & Bowman, D.M.J.S. (2015b). Macroecology of Australian tall eucalypt forests: Baseline data from a continental-scale permanent plot network. *PLoS ONE, 10*, 1-24

Woodcock, C.E., & Strahler, A.H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment, 21*, 311-332

Woodgate, W., Jones, S.D., Suarez, L., Hill, M.J., Armston, J.D., Wilkes, P., Soto-Berelov, M., Haywood, A., & Mellor, A. (2015). Understanding the variability in ground-based methods for retrieving canopy openness, gap fraction, and leaf area index in diverse forest systems. *Agricultural and Forest Meteorology, 205*, 83-95

Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., & White, J.C. (2008). The role of LiDAR in sustainable forest management. *Forestry Chronicle, 84*, 807-826

Xia, J., Broadhurst, D.I., Wilson, M., & Wishart, D.S. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics, 9*, 280-299

Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology, 1*, 9-23

Xu, X., Du, H., Zhou, G., Ge, H., Shi, Y., Zhou, Y., Fan, W., & Fan, W. (2011). Estimation of aboveground carbon stock of Moso bamboo (Phyllostachys heterocycla var. pubescens) forest with a Landsat Thematic Mapper image AU - Xu, Xiaojun. *International Journal of Remote Sensing, 32*, 1431-1448

Yang, G., Zhao, Y., Li, B., Ma, Y., Li, R., Jing, J., & Dian, Y. (2019). Tree species classification by employing multiple features acquired from integrated sensors. *Journal of Sensors, 2019*, 12

Yang, R., Su, L., Zhao, X., Wan, H., & Sun, J. (2017a). Representative band selection for hyperspectral image classification. *Journal of Visual Communication and Image Representation, 48*, 396-403

Yang, Y., Qian, W., & Zou, H. (2017b). Insurance premium prediction via gradient tree-boosted tweedie compound poisson models. *Journal of Business & Economic Statistics*, 1-15

Yang, Z., Willis, P., & Mueller, R. (2008). Impact of band-ratio enhanced AWIFS image on crop classification accuracy. In, *Pecora 17 – The Future of Land Imaging…Going Operational* (p. 11). Denver, Colorado

Yee, M. (2005). The ecology and habitat requirements of saproxylic beetles native to Tasmanian wet eucalypt forests: potential impacts of commercial forestry practices. In, *School of Agricultural Science* (p. 203). Hobart, Tasmania, Australia: PhD Thesis. University of Tasmania

Youngentob, K.N., Roberts, D.A., Held, A.A., Dennison, P.E., Jia, X., & Lindenmayer, D.B. (2011). Mapping two *Eucalyptus* subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data. *Remote Sensing of Environment, 115*, 1115-1128

Yuan, Z., Zhao, Y., Dan, Z., & Bingfang, W. (2016). Forest biodiversity mapping using airborne LiDAR and hyperspectral data. In, *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*. No.7 Tianchen East Road, Chaoyang District, Beijing 100105, China: IEEE

Zahedi, S., Shahedi, K., Habibnejad Rawshan, M., Solimani, K., & Dadkhah, K. (2017). Soil depth modelling using terrain analysis and satellite imagery: the case study of Qeshlaq mountainous watershed (Kurdistan, Iran). *Journal of Agricultural Engineering, 48*, 167-174

Zald, H.S.J., Ohmann, J.L., Roberts, H.M., Gregory, M.J., Henderson, E.B., McGaughey, R.J., & Braaten, J. (2014). Influence of LiDAR, Landsat imagery, disturbance history, plot location accuracy, and plot size on accuracy of imputation maps of forest composition and structure. *Remote Sensing of Environment, 143*, 26-38

Zald, H.S.J., Wulder, M.A., White, J.C., Hilker, T., Hermosilla, T., Hobart, G.W., & Coops, N.C. (2016). Integrating Landsat pixel composites and change metrics with LiDAR plots to predictively map forest structure and aboveground biomass in Saskatchewan, Canada. *Remote Sensing of Environment, 176*, 188-201

Zehm, A., Nobis, M., & Schwabe, A. (2003). Multiparameter analysis of vertical vegetation structure based on digital image processing. *Flora - Morphology, Distribution, Functional Ecology of Plants, 198*, 142-160

Zellweger, F., Morsdorf, F., Purves, R.S., Braunisch, V., & Bollmann, K. (2014). Improved methods for measuring forest landscape structure: LiDAR complements field-based habitat assessment. *Biodiversity and Conservation, 23*, 289-307

Zhang, C. (2014). Combining hyperspectral and LiDAR data for vegetation mapping in the Florida everglades. *Photogrammetric Engineering and Remote Sensing, 80*, 733-743

Zhang, C. (2015). Applying data fusion techniques for benthic habitat mapping and monitoring in a coral reef ecosystem. *ISPRS Journal of Photogrammetry and Remote Sensing, 104*, 213-223

Zhang, C., & Qiu, F. (2012). Mapping individual tree species in an urban forest using airborne LiDAR data and hyperspectral imagery. *Photogrammetric Engineering & Remote Sensing, 78*, 1079-1087

Zhang, C., Selch, D., Xie, Z., Roberts, C., Cooper, H., & Chen, G. (2013a). Object-based benthic habitat mapping in the Florida keys from hyperspectral imagery. *Estuarine, Coastal and Shelf Science, 134*, 88-97

Zhang, C., & Xie, Z. (2013). Object-based vegetation mapping in the Kissimmee river watershed using Hymap data and machine learning techniques. *Wetlands, 33*, 233-244

Zhang, C., Xie, Z., & Selch, D. (2013b). Fusing LiDAR and digital aerial photography for object-based forest mapping in the Florida Everglades. *GIScience and Remote Sensing, 50*, 562-573

Zhang, N. (2007). Scale issues in ecology: Upscaling. *Acta Ecologica Sinica, 27*, 4252-4266

Zhang, N., Hong, Y., Qin, Q., & Zhu, L. (2013c). Evaluation of the visible and shortwave infrared drought index in China. *International Journal of Disaster Risk Science, 4*, 68-76

Zhang, W., & Montgomery, D.R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water resources research, 30*, 1019-1028

Zhang, Z., Kazakova, A., Moskal, L.M., & Styers, D.M. (2016). Object-based tree species classification in urban ecosystems using LiDAR and hyperspectral data. *Forests, 7*

Zhang, Z., Liu, X., Peterson, J., & Wright, W. (2011). Cool temperate rainforest and adjacent forests classification using airborne LiDAR data. *Area, 43*, 438-448

Zhang, Z., Liu, X., & Wright, W. (2014). Feature assessment in object-based forest classification using airborne LiDAR data and high spatial resolution satellite imagery. In Q. Weng, P. Gamba, G. Xian, G. Wang, & J. Zhu (Eds.), *3rd International Workshop on Earth Observation and Remote Sensing Applications, EORSA 2014 - Proceedings* (pp. 393-397). Changsha, China: IEEE

Zhao, Y., Zeng, Y., Zhao, D., Wu, B., & Zhao, Q. (2016). The optimal leaf biochemical selection for mapping species diversity based on imaging spectroscopy. *Remote Sensing, 8*, 216

Zhao, Y., Zeng, Y., Zheng, Z., Dong, W., Zhao, D., Wu, B., & Zhao, Q. (2018). Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. *Remote Sensing of Environment, 213*, 104-114

Zhou, Y., & Qiu, F. (2015). Fusion of high spatial resolution WorldView-2 imagery and LiDAR pseudo-waveform for object-based image analysis. *ISPRS Journal of Photogrammetry and Remote Sensing, 101*, 221-232

Ziadat, F.M. (2005). Analyzing digital terrain attributes to predict soil attributes for a relatively large area. *Soil Science Society of America Journal, 69*, 1590-1599

Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., & Gerard, P.D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment, 87*, 171-182

Zlinszky, A., Heilmeier, H., Balzter, H., Czúcz, B., & Pfeifer, N. (2015). Remote sensing and GIS for habitat quality monitoring: New approaches and future research. *Remote Sensing*, *7*, 7987-7994